

# A fitting formula for the non-Gaussian contribution to the lensing power spectrum covariance

J. Pielorz, J. Rödiger, I. Tereno, and P. Schneider

Argelander-Institut für Astronomie (AIfA), Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany  
e-mail: pielorz@astro.uni-bonn.de

Received; accepted

## ABSTRACT

**Context.** Weak gravitational lensing is one of the most promising tools to investigate the equation-of-state of dark energy. In order to obtain reliable parameter estimations for current and future experiments, a good theoretical understanding of dark matter clustering is essential. Of particular interest is the statistical precision to which weak lensing observables, such as cosmic shear correlation functions, can be determined.

**Aims.** We construct a fitting formula for the non-Gaussian part of the covariance of the lensing power spectrum. The Gaussian contribution to the covariance, which is proportional to the lensing power spectrum squared, and optionally shape noise can be included easily by adding their contributions.

**Methods.** Starting from a canonical estimator for the dimensionless lensing power spectrum, we model first the covariance in the halo model approach including all four halo terms for one fiducial cosmology and then fit two polynomials to the expression found. On large scales, we use a first-order polynomial in the wave-numbers and dimensionless power spectra that goes asymptotically towards  $1.1 C_{\text{pt}}$  for  $\ell \rightarrow 0$ , i.e., the result for the non-Gaussian part of the covariance using tree-level perturbation theory. On the other hand, for small scales we employ a second-order polynomial in the dimensionless power spectra for the fit.

**Results.** We obtain a fitting formula for the non-Gaussian contribution of the convergence power spectrum covariance that is accurate to 10% for the off-diagonal elements, and to 5% for the diagonal elements, in the range  $50 \lesssim \ell \lesssim 5000$  and can be used for single source redshifts  $z_s \in [0.5, 2.0]$  in WMAP5-like cosmologies.

**Key words.** gravitational lensing – Methods: *N*-body simulations – Cosmology: theory – large-scale structure of the Universe

## 1. Introduction

Weak gravitational lensing by the large-scale structure, or cosmic shear, is an important tool to probe the mass distribution in the Universe and to estimate cosmological parameters. The constraints it provides are independent and complementary to those found by other cosmological probes such as cosmic microwave background (CMB) anisotropies, supernovae (SN) type Ia, baryon acoustic oscillations (BAO) or galaxy redshift surveys. The cosmic shear field quantifies the distortion of faint galaxy images that is induced by continuous light deflections caused by the large-scale structure in the Universe (e.g., Bartelmann & Schneider 2001; Schneider 2006). Since this effect is too small to be measured for a single galaxy, large surveys with millions of galaxies are required to detect it in a statistical way. The cosmic shear signal has been successfully measured in various surveys, since the first detections of Bacon et al. (2000); Kaiser et al. (2000); Van Waerbeke et al. (2000); Wittman et al. (2000). Most recently, shear two-point correlation functions were measured in the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS) and were used to constrain the amplitude of dark matter clustering,  $\sigma_8 \Omega_m^{0.6}$ , with 5% uncertainty (Fu et al. 2008).

The next generation of galaxy surveys will greatly improve the precision with which weak lensing effects can be measured (Albrecht et al. 2006) enabling us to obtain, with accurate redshift information and tomographic measurements, precise constraints on the evolution of dark energy. However, the expected improvement in future data only leads to a significant improvement of the precision and accuracy of the cosmological interpretation if the systematic errors, the underlying physics, and the statistical precision of cosmic shear estimators are well understood. Systematics currently identified arise mainly from non-cosmological sources of shear correlations, i.e., intrinsic alignments of galaxies (e.g., Schaefer 2008, for a review), and biases on the shear measurement (Massey et al. 2007; Semboloni et al. 2008). This paper addresses the issue of the statistical precision of cosmic shear estimators, determined by the covariance of the estimator. Since much of the scales probed by cosmic shear lie in the non-linear regime, being affected by non-linear clustering, the covariance depends on non-Gaussian effects and has a non-Gaussian, as well as a Gaussian, contribution. Indeed, even though the non-Gaussianity of the shear field is weaker than that of the matter field due to the projection along the line-of-sight, various studies indicate that the non-Gaussian contribution to the covariance cannot be neglected when constraining cosmological parameters with weak lensing (Scoccimarro et al. 1999; White & Hu 2000; Cooray & Hu 2001; Kilbinger & Schneider 2005; Semboloni et al. 2007; Takada & Jain 2009).

Most cosmic shear results are based on the measurement of two-point correlation functions of the shear field. Since, in general, the number of independent measurements is insufficient to infer the complete covariance directly from observations, one may derive it from ray-tracing maps of numerical *N*-body simulations. This, however, requires a large number of realizations and, in addition, is very time-consuming if an exploration of the covariance in the parameter space is needed. An alternative is to compute

the covariance with an analytic approach. For shear two-point correlation functions, Schneider et al. (2002) derived an expression for the Gaussian contribution to the covariance. Semboloni et al. (2007) fitted the ratio between that expression and a covariance computed with  $N$ -body simulations, containing both Gaussian and non-Gaussian contributions, providing thus a formula to compute the total covariance from the Gaussian term. In Fourier space, large-scale modes are independent and, differently from real space, the Gaussian contribution to the covariance of the convergence power spectrum (i.e., of the Fourier transform of the two-point shear correlation function) is diagonal and can be computed from the convergence power spectrum alone (Kaiser 1992; Joachimi et al. 2008), whereas the non-Gaussian contribution can be computed from the trispectrum of the convergence (Scoccimarro et al. 1999). The trispectrum, on large scales, can be accurately derived in tree-level perturbation theory<sup>1</sup>, and, on small scales, is well represented by the one-halo term of a halo model approach. A non-Gaussian part of the covariance consisting of a perturbation theory term and a one-halo term was used, e.g., in Takada & Jain (2009).

This paper aims at producing an accurate expression for the non-Gaussian contribution of the covariance of the convergence power spectrum that is fast to compute, contributing thus to accurate estimates of cosmological parameters. Following Scoccimarro et al. (1999) and Cooray & Hu (2001), we start from a canonical estimator of the dimensionless convergence power spectrum and use it to derive an analytic expression for the corresponding covariance. The various spectra involved are evaluated using the halo model approach of dark matter clustering (Seljak 2000; Ma & Fry 2000; Scoccimarro et al. 2001; Cooray & Sheth 2002). The halo model approach assumes that all dark matter in the Universe is bound in spherical halos, and uses results from numerical  $N$ -body simulations to characterize halo properties such as their profile, abundance and clustering behavior.

The evaluation of the covariance of the convergence power spectrum in the halo model approach is time-consuming. In addition, it may be needed to repeat it for different cosmological models for the purpose of parameter estimation. To allow for a faster computation, we construct a fitting formula for the non-Gaussian part of the convergence power spectrum covariance. On small scales, we fit the halo model result with a polynomial in the non-linear dimensionless convergence power spectrum. On large scales, we fit the ratio between the halo model covariance and the perturbation theory covariance. We stress that it is a fit to the halo model covariance, not involving a covariance computed from  $N$ -body simulations. The result is, however, calibrated by  $N$ -body simulations, since they determine the halo model parameters.

The paper is organized as follows. We define in Sect. 2 the reference cosmology, considering the growth of matter perturbations. We introduce the convergence spectra, construct an estimator for the dimensionless convergence power spectrum, and derive an expression for its covariance in Sect. 3. In Sect. 4, we describe the halo model approach, and compute the covariance of the power spectrum. The covariance depends on the values of halo model parameters, which are also defined here. It also depends on the power spectrum, bispectrum and trispectrum of the correlations of halo centers. Expressions for these spectra, in the framework of perturbation theory, are given in the Appendix. Section 5 tests the accuracy of the halo model predictions, for both the power spectrum and its covariance, against two sets of ray-tracing simulations. Section 6 presents the fitting formula for the non-Gaussian contribution to the covariance where its coefficients are given as function of source redshift. We conclude in Sect. 7.

## 2. Structure formation in a $\Lambda$ CDM cosmology

Throughout this work we assume a spatially flat cold dark matter model with a cosmological constant ( $\Omega_m + \Omega_\Lambda = 1$ ), as supported by the latest 5-year data release of WMAP results (Komatsu et al. 2009). The expansion rate of the Universe,  $H(a) \equiv \dot{a}/a$ , in such models is described by the Friedmann equation  $H^2(a) = H_0^2 (\Omega_m a^{-3} + \Omega_\Lambda)$ , where  $H_0 \equiv 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}$  is the Hubble constant,  $\Omega_m$  denotes the combined contributions from dark matter and baryons today in terms of the critical density  $\rho_{\text{crit}} \equiv 3H_0^2/(8\pi G)$ , and  $\Omega_\Lambda$  is the density parameter of the cosmological constant. The comoving distance to a source at  $a$  is then

$$w(a) = \int_a^1 \frac{c da'}{a'^2 H(a')}, \quad (1)$$

where the scale factor is related to the redshift via the relation  $1 + z = 1/a$  using the convention  $a(t_0) = 1$  today.

In structure formation, the central quantity is the Fourier transform of the density contrast  $\delta(\mathbf{x}, t) = [\rho(\mathbf{x}, t) - \bar{\rho}(t)]/\bar{\rho}(t)$ , which describes the relative deviation of the local matter density  $\rho(\mathbf{x}, t)$  to the comoving average density of the Universe  $\bar{\rho}(t)$  at time  $t$ . We suppress the time dependence of  $\delta$  in the following. In this way, the mean density contrast is by definition zero, and we can describe matter perturbations in the early Universe as zero-mean Gaussian random fields. In this case, the statistical properties of the Fourier transformed density field,

$$\tilde{\delta}(\mathbf{k}) = \int d^3x e^{i\mathbf{k}\cdot\mathbf{x}} \delta(\mathbf{x}), \quad (2)$$

are completely characterized by the power spectrum

$$\langle \tilde{\delta}(\mathbf{k}_1) \tilde{\delta}(\mathbf{k}_2) \rangle \equiv (2\pi)^3 \delta_D(\mathbf{k}_1 + \mathbf{k}_2) P_\delta(k_1), \quad (3)$$

where  $\langle \cdot \rangle$  is the ensemble average and  $\delta_D$  denotes the Dirac delta distribution. Note that throughout this paper the tilde symbol is used to denote the Fourier transform of the corresponding quantity.

In linear perturbation theory, which is valid on large scales, the power spectrum at a scale factor  $a$  is characterized by

$$P_{\text{lin}}(k, a) = A k^{n_s} T^2(k) D^2(a), \quad (4)$$

<sup>1</sup> We refer by tree-level perturbation theory to the lowest, non-vanishing order of the considered quantity in perturbation theory.

where the amplitude  $A$  is normalized in terms of  $\sigma_8$ ,  $n_s$  denotes the spectral index of the primordial power spectrum, and  $T(k)$  is the transfer function. Note that all Fourier modes of the matter density grow at the same rate, i.e.,  $\tilde{\delta}(\mathbf{k}, a) = \tilde{\delta}(\mathbf{k})D(a)$ , where

$$D(a) \propto \frac{H(a)}{H_0} \int_0^a \frac{da'}{[a' H(a')/H_0]^3} \quad (5)$$

is the growth factor which we normalize as  $D(a = 1) = 1$ . In the non-linear regime, i.e., on small scales, different Fourier modes couple and the Gaussian assumption cannot be maintained. Thus we have to consider higher-order moments of the density field to describe its statistical properties. In perturbation theory, it is possible to find analytic expressions for these moments, which hold up to the quasi-linear regime. In Appendix A, we derive the expressions for the bispectrum and trispectrum in tree-level perturbation theory, which are the Fourier transforms of the three- and four-point-correlation functions, respectively.

### 3. Covariance of the convergence power spectrum

A central quantity in weak lensing applications is the two-dimensional projection of the density contrast  $\delta(w\boldsymbol{\theta}, w)$  on the sky, which is known as *effective convergence*  $\kappa(\boldsymbol{\theta})$ . It is obtained by projecting the density contrast along the backward-directed light-cone of the observer according to

$$\kappa(\boldsymbol{\theta}) = \int_0^{w_H} dw w G(w) \delta(w\boldsymbol{\theta}, w), \quad (6)$$

where  $w \equiv w(z)$  denotes the redshift-dependent comoving distance,  $w_H$  is the comoving distance to the horizon and the weight function  $G(w)$  takes into account the distribution of source galaxies along the line-of-sight. We assume for simplicity that all background sources are situated at a single comoving distance  $w_s \equiv w(z_s)$ , such that the weight function has the form

$$G(w) = \frac{3}{2} \Omega_m \left( \frac{H_0}{c} \right)^2 a^{-1} \frac{w_s - w}{w_s} H(w_s - w), \quad (7)$$

where  $H(x)$  denotes the Heaviside step function. To take advantage of the Fourier properties, we analyze the statistical properties of the Fourier counterpart of  $\kappa(\boldsymbol{\theta})$ . For the theoretical consideration of the convergence power spectrum covariance we need the second- and the fourth-order moments, as will become apparent later. They are defined by

$$\langle \tilde{\kappa}(\boldsymbol{\ell}_1) \tilde{\kappa}(\boldsymbol{\ell}_2) \rangle \equiv (2\pi)^2 \delta_D(\boldsymbol{\ell}_{12}) P_\kappa(\boldsymbol{\ell}_1), \quad (8)$$

$$\langle \tilde{\kappa}(\boldsymbol{\ell}_1) \tilde{\kappa}(\boldsymbol{\ell}_2) \tilde{\kappa}(\boldsymbol{\ell}_3) \tilde{\kappa}(\boldsymbol{\ell}_4) \rangle_c \equiv (2\pi)^2 \delta_D(\boldsymbol{\ell}_{1234}) T_\kappa(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3, \boldsymbol{\ell}_4), \quad (9)$$

where the subscript ‘c’ refers to the connected part of the corresponding moment and  $\boldsymbol{\ell}_{i\dots j} = \boldsymbol{\ell}_i + \dots + \boldsymbol{\ell}_j$  is a sum of Fourier wave-vectors. The convergence power spectrum and trispectrum are calculated using the flat-sky and Limber’s approximation (Kaiser 1998; Scoccimarro et al. 1999; Bernardeau et al. 2002):

$$P_\kappa(\boldsymbol{\ell}) = \int_0^{w_H} dw G^2(w) P_\delta\left(\frac{\boldsymbol{\ell}}{w}, w\right), \quad (10)$$

$$T_\kappa(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3, \boldsymbol{\ell}_4) = \int_0^{w_H} dw \frac{G^4(w)}{w^2} T_\delta\left(\frac{\boldsymbol{\ell}_1}{w}, \frac{\boldsymbol{\ell}_2}{w}, \frac{\boldsymbol{\ell}_3}{w}, \frac{\boldsymbol{\ell}_4}{w}, w\right), \quad (11)$$

where  $P_\delta$  and  $T_\delta$  are the corresponding three-dimensional matter power spectrum and trispectrum (Fourier transform of the four-point correlation function).

We are interested in estimating the dimensionless convergence power spectrum

$$\mathcal{P}_\kappa(\ell) = (\ell^2/2\pi) P_\kappa(\ell), \quad (12)$$

and the corresponding covariance for wave-vectors of different length  $\ell$ . A natural choice for the estimator of the dimensionless convergence power spectrum is (Scoccimarro et al. 1999; Cooray & Hu 2001; Takada & Bridle 2007)

$$\hat{\mathcal{P}}_\kappa(\ell_i) = \frac{1}{A} \int_{\ell \in \ell_i} \frac{d^2\ell}{A_r(\ell_i)} \frac{\ell^2}{2\pi} \tilde{\kappa}(\boldsymbol{\ell}) \tilde{\kappa}(-\boldsymbol{\ell}), \quad (13)$$

which is unbiased in the limit of infinitesimal small bin sizes, since  $\langle \hat{\mathcal{P}}_\kappa(\ell_i) \rangle \equiv \mathcal{P}_\kappa(\ell_i)$ . Here,  $A = 4\pi f_{\text{sky}}$  denotes the solid angle of a survey with a fractional sky coverage of  $f_{\text{sky}}$ , and the integration is performed over the Fourier modes lying in the annulus defined by  $\ell_i - \Delta\ell_i/2 \leq \ell \leq \ell_i + \Delta\ell_i/2$ , where  $\Delta\ell_i$  is the width of the  $i$ -th bin. We denote the integration area formed by the annulus as  $A_r(\ell_i)$ .

The evaluation of the covariance of the estimator, Eq. (13), results in an expression of the form

$$\begin{aligned} C_{ij} &\equiv \text{Cov}[\hat{\mathcal{P}}_\kappa(\ell_i), \hat{\mathcal{P}}_\kappa(\ell_j)] = \langle \hat{\mathcal{P}}_\kappa(\ell_i) \hat{\mathcal{P}}_\kappa(\ell_j) \rangle - \langle \hat{\mathcal{P}}_\kappa(\ell_i) \rangle \langle \hat{\mathcal{P}}_\kappa(\ell_j) \rangle \\ &= \frac{1}{A} \left[ \frac{(2\pi)^2}{A_r(\ell_i)} 2\mathcal{P}_\kappa^2(\ell_i) \delta_{\ell_i \ell_j} + \bar{T}_\kappa(\ell_i, \ell_j) \right] \equiv C_{ij}^G + C_{ij}^{\text{NG}}, \end{aligned} \quad (14)$$

where  $\mathcal{P}_k(\ell_i)$  is given by Eq. (12), and  $\delta_{\ell_i \ell_j}$  denotes the Kronecker delta. The second term in square brackets is the bin-averaged convergence trispectrum,

$$\bar{T}_k(\ell_i, \ell_j) = \int_{|\ell_1| \in \ell_i} \frac{d^2 \ell_1}{A_r(\ell_i)} \int_{|\ell_2| \in \ell_j} \frac{d^2 \ell_2}{A_r(\ell_j)} \frac{\ell_1^2 \ell_2^2}{(2\pi)^2} T_k(\ell_1, -\ell_1, \ell_2, -\ell_2), \quad (15)$$

where  $T_k$  is the non-linear convergence trispectrum as defined in Eq. (11). To derive Eq. (14), we made use of the definitions of the convergence power spectrum and trispectrum in Eqs. (8) and (9), and of the discrete limit of the delta distribution  $\delta_D(\mathbf{0}) \rightarrow A/(2\pi)^2$ . The derived expression consists of two terms: a Gaussian part  $C^G$ , which scales as the convergence power spectrum squared and only contributes to the diagonal of the covariance matrix (see, e.g. Joachimi et al. 2008), and a non-Gaussian part  $C^{NG}$ , which scales as the dimensionless bin-averaged convergence trispectrum and introduces correlations between the wave-vectors of different bins (see, e.g. Scoccimarro et al. 1999; Cooray & Hu 2001). Both terms are inversely proportional to the survey area  $A$ , but have a different behavior with respect to the bin width  $\Delta \ell_i$ . While the Gaussian term decreases with increasing bin size, the non-Gaussian term is independent of the binning, since the bin area cancels out after the integration.

We can analytically perform one of the integrations of the bin-averaged trispectrum in Eq. (15). First, note that it only depends on the parallelogram configuration of the convergence trispectrum, i.e., setting  $\ell_2 = -\ell_1$ ,  $\ell_3 = \ell_2$  and  $\ell_4 = -\ell_2$  in Eq. (11). Also, if we choose an appropriate coordinate system for the integration over the wave-vectors, the problem becomes symmetric under rotations and we can parametrize the convergence trispectrum by the length of the two sides of the parallelogram  $\ell_1$  and  $\ell_2$  and the angle between them,  $\cos \varphi = (\ell_1 \cdot \ell_2)/\ell_1 \ell_2$ . Hence, we define

$$T_k(\ell_1, \ell_2, \cos \varphi) \equiv T_k(\ell_1, -\ell_1, \ell_2, -\ell_2). \quad (16)$$

Making use of the symmetry properties of this problem, one angular integration becomes trivial and the integration in Eq. (15) simplifies to

$$\bar{T}_k(\ell_i, \ell_j) = \frac{1}{2\pi} \int_{|\ell_1| \in \ell_i} \frac{d\ell_1}{A_r(\ell_i)} \ell_1^3 \int_{|\ell_2| \in \ell_j} \frac{d\ell_2}{A_r(\ell_j)} \ell_2^3 \int_0^{2\pi} d\varphi T_k(\ell_1, \ell_2, \cos \varphi). \quad (17)$$

If the bin-width  $\Delta \ell_i$  is sufficiently small ( $\Delta \ell_i \ll \ell_i$ ), the integration area is  $A_r(\ell_i) = 2\pi \ell_i \Delta \ell_i$ , and we can make use of the mean value theorem. In this way, we approximate the integral in Eq. (17) by an angular average

$$\bar{T}_k(\ell_i, \ell_j) \simeq \frac{1}{2\pi} \int_0^{2\pi} d\varphi \frac{\ell_i^2 \ell_j^2}{(2\pi)^2} T_k(\ell_i, \ell_j, \cos \varphi). \quad (18)$$

Note that if  $T_k(\ell_i, \ell_j, \cos \varphi)$  is independent of the angle between  $\ell_i$  and  $\ell_j$ , an approximation of the covariance can be calculated without having to perform an integration at all. In particular, this is the case for the 1-halo term of the three-dimensional matter trispectrum, as we will see later in Eq. (42).

## 4. Halo Model

We have seen that the covariance of the dimensionless convergence power spectrum estimator consists of two terms: a Gaussian part, which is proportional to the dimensionless convergence power spectrum squared and a non-Gaussian part, which is the bin-averaged dimensionless convergence trispectrum (see Eq. 14). We will compute these terms using the halo model approach (Seljak 2000; Ma & Fry 2000; Scoccimarro et al. 2001; see also the comprehensive review by Cooray & Sheth 2002).

### 4.1. Overview

With the assumption that all dark matter is bound in spherically-symmetric, virialized halos, the halo model provides a way to calculate the three-dimensional polyspectra of dark matter in the non-linear regime. In Sect. 4.5 below, we summarize the equations one obtains for the dark matter power spectrum and trispectrum.

In the halo model description, the density field at an arbitrary position  $\mathbf{x}$  in space is given as a superposition of all  $N$  halo density profiles such that

$$\rho(\mathbf{x}) = \sum_{i=1}^N f(\mathbf{x} - \mathbf{x}_i; m_i, c_i) \equiv \sum_{i=1}^N m_i u(\mathbf{x} - \mathbf{x}_i; m_i, c_i), \quad (19)$$

where  $f(\mathbf{x} - \mathbf{x}_i; m_i, c_i)$  denotes the density profile of the  $i$ -th halo with center of mass at  $\mathbf{x}_i$  and  $u \equiv f/m_i$  is the normalized profile. By parametrizing the halo profile in this way, we assume that the shape of the  $i$ -th halo depends only on the halo mass  $m_i$  and the halo concentration parameter  $c_i$ , which we define below. The dark matter polyspectra of the density field  $\rho(\mathbf{x})$  follow then from taking the ensemble averages,  $\langle X \rangle$ , of products of the density at different points in space. Assuming that the number of halos is  $N = \bar{n}V$ , where  $\bar{n}$  is the average number density of halos and  $V$  the considered volume, we compute the ensemble averages by integrating over the joint probability density function (PDF) for the  $N$  halos that form the field (Smith & Watts 2005), i.e.,

$$\langle X \rangle \equiv \int \left[ \prod_{i=1}^N d^3 x_i dm_i dc_i \right] p(\mathbf{x}_1, \dots, \mathbf{x}_N, m_1, \dots, m_N, c_1, \dots, c_N) X, \quad (20)$$

where  $p(\cdot, \cdot, \cdot)$  denotes the PDF. If one considers that position and mass of a single halo are independent random variables, the PDF factorizes as

$$p(\mathbf{x}; m, c) = p(\mathbf{x})p(m)p(c|m) = \frac{1}{V} \frac{n(m)}{\bar{n}} p(c|m), \quad (21)$$

where  $n(m)$  is the halo mass function and  $p(c|m)$  is the concentration probability distribution for halos given a mass  $m$ .

#### 4.2. Ingredients

The halo model approach provides a scale-dependent description of the statistical properties of the large-scale structure. On small scales, the correlation of dark matter is governed by the mass profiles of the halos, whereas on large scales the clustering between different halos determines the nature of the correlation. As there are a multitude of models to describe the behavior on different scales, and an even larger number of parameters one has to set judiciously, there exists no such thing as a unique halo model. In order to have reproducible results, it is therefore necessary to specify one's choice of parameters. For this work, we will adopt the following parameters for the halo model:

1. The average mass of a halo is defined as the mass within a sphere of virial radius  $r_{\text{vir}}$  as  $m \equiv (4\pi/3)r_{\text{vir}}^3 \Delta_{\text{vir}} \bar{\rho}$ , where  $\Delta_{\text{vir}}$  denotes the overdensity of the virialized halo with respect to the average comoving mass density  $\bar{\rho}$  in the Universe. Typically, values for  $\Delta_{\text{vir}}$  are derived in the framework of the non-linear spherical collapse model (e.g. Gunn & Gott 1972). Expressions valid for different cosmologies are summarized in Nakamura & Suto (1997). In our implementation, we use the results which are valid for a flat  $\Lambda$ CDM-Universe, i.e.,

$$\Delta_{\text{vir}}(z) = 18\pi^2(1 + 0.4093x^{2.71572}), \quad (22)$$

where  $x \equiv (\Omega_m^{-1} - 1)^{1/3}/(1+z)$ . We find for our fiducial WMAP5-like cosmology  $\Delta_{\text{vir}}(z=0) = 349$ .

2.  $N$ -body simulations suggest that the density profile of a halo follows a universal function. We choose to use the NFW profile (Navarro et al. 1997), which is in good agreement with numerical results and has an analytical Fourier transform. It is given by

$$\rho(r, m) = \frac{\rho_s}{(r/r_s)(1 + r/r_s)^2}, \quad (23)$$

where  $\rho_s$  is the amplitude of the density profile and  $r_s$  characterizes the scale at which the slope of the density profile changes. For small scales ( $r \lesssim r_s$ ) the profile scales with  $\rho \propto r^{-1}$ , whereas for large scales it behaves as  $\rho \propto r^{-3}$ . The Fourier transform of the NFW profile is

$$\begin{aligned} \tilde{u}(k; m, c) &= \int d^3x \rho(\mathbf{x}; m, c) e^{i\mathbf{k}\cdot\mathbf{x}} \bigg/ \int d^3x \rho(\mathbf{x}; m, c) = \int_0^{r_{\text{vir}}} dr 4\pi r^2 \frac{\sin(kr)}{kr} \frac{\rho(r; m, c)}{m} \\ &= \left[ \ln(1+c) - \frac{c}{1+c} \right]^{-1} \left\{ \sin \eta [\text{Si}([1+c]\eta) - \text{Si}(\eta)] + \cos \eta [\text{Ci}([1+c]\eta) - \text{Ci}(\eta)] - \frac{\sin(c\eta)}{(1+c)\eta} \right\}, \end{aligned} \quad (24)$$

where  $\eta \equiv kr_{\text{vir}}/c$ , we truncated the integration at  $r_{\text{vir}}$  in the second step, and introduced the concentration parameter  $c \equiv r_{\text{vir}}/r_s$  in the third step. Additionally, we use for the sine- and cosine integrals the definitions

$$\text{Ci}(x) = - \int_x^\infty \frac{\cos t}{t} dt, \quad \text{Si}(x) = \int_0^x \frac{\sin t}{t} dt. \quad (25)$$

3. The abundance of halos of mass  $m$  at a redshift  $z$  is given by

$$n(m, z) = \frac{\bar{\rho}}{m^2} \frac{d \ln \nu}{d \ln m} \nu f(\nu), \quad (26)$$

where we introduced the dimensionless variable  $\nu = \nu(m, z)$ ,

$$\nu(m, z) \equiv \frac{\delta_{\text{sc}}(z)}{D(z)\sigma(m)}, \quad (27)$$

where  $D(z)$  denotes the redshift-dependent growth factor,  $\sigma^2(m)$  is the smoothed variance of the density contrast, and  $\delta_{\text{sc}}(z)$  denotes the value of a spherical overdensity that collapses at a redshift  $z$  as calculated from linear perturbation theory. In our work, we use the expression from Nakamura & Suto (1997), which is valid for a  $\Lambda$ CDM Universe,

$$\delta_{\text{sc}}(z) = \frac{3(12\pi)^{2/3}}{20} [1 - 0.0123 \ln(1+x^3)], \quad (28)$$

where  $x \equiv (\Omega_m^{-1} - 1)^{1/3}/(1+z)$ . The quantity has only a weak dependence on redshift and we find  $\delta_{\text{sc}}(z=0) = 1.675$  for our fiducial model.

The advantage of introducing  $\nu$  is that part of the mass function can be expressed by the multiplicity function  $\nu f(\nu)$ , which has a universal shape, i.e., is independent of cosmological parameters and redshift. In this work, we employ the Sheth and Tormen mass function (Sheth & Tormen 1999)

$$\nu f(\nu) = A(p) [1 + (q\nu^2)^{-p}] \sqrt{\frac{2q}{\pi}} \nu \exp(-q\nu^2/2), \quad (29)$$

which is an improvement over the original Press-Schechter formulation (Press & Schechter 1974). We use the parameter values  $p = 0.3$ ,  $q = 0.707$ , and amplitude  $A(0.3) = 0.322$ , which follows from mass conservation.

4. The concentration parameter  $c \equiv r_{\text{vir}}/r_s$  characterizes the form of the halo profile. From  $N$ -body simulations one finds that the average,  $\bar{c}$ , depends on the halo mass (Bullock et al. 2001) like

$$\bar{c}(m, z) = \frac{c_*}{1+z} \left( \frac{m}{m_*} \right)^{-\alpha}, \quad (30)$$

where  $m_* = m_*(z=0)$  is the characteristic mass defined within the Press-Schechter formalism as  $\delta_{\text{sc}}(z=0) = \sigma(m_*)$ . In the following, we will use the values  $c_* = 10$  and  $\alpha = 0.2$  as proposed by Takada & Jain (2003). This implies that more massive halos are less centrally concentrated than less massive ones. However, results from numerical  $N$ -body simulations (Jing 2000; Bullock et al. 2001) indicate that there is a significant scatter in the concentration parameter for halos of the same mass. Furthermore, Jing (2000) proposes that such a concentration distribution can be described by a log-normal distribution

$$p(c|m)dc = \frac{1}{\sqrt{2\pi\sigma_{\ln c}^2}} \exp \left[ -\frac{(\ln c - \ln \bar{c})^2}{2\sigma_{\ln c}^2} \right] d \ln c. \quad (31)$$

Typical values for the concentration dispersion range from  $\sigma_{\ln c} = 0.18$  to  $\sigma_{\ln c} = 0.32$  (Jing 2000; Wechsler et al. 2002). Note that the width of the distribution  $\sigma_{\ln c}$  is independent of the halo mass. The variation of the halo concentration can be attributed to the different merger histories of the halos (Wechsler et al. 2002). We will analyze the impact of this effect on different spectra in Sect. 4.7. When we use only the mean concentration parameter, we have to replace the probability distribution of the concentration, needed for example in Eq. (21), by a Dirac delta distribution

$$p(c|m)dc = \delta_D(c - \bar{c}) c d \ln c. \quad (32)$$

5. On large scales, the correlation of the dark matter density field is governed by the spatial distribution of halos. Since the clustering behavior of halos and matter density differ, one introduces the bias factors  $b_i(m, z)$  such that

$$\delta_h(\delta) \equiv \delta_h(\mathbf{x}; m, z) = b_1(m, z)\delta(\mathbf{x}) + \frac{b_2(m, z)}{2}\delta^2(\mathbf{x}) + \dots \quad (33)$$

In this way, the halo density contrast,  $\delta_h(\delta)$ , is expressed as a Taylor expansion of the matter density contrast,  $\delta(\mathbf{x})$ . The bias parameters are in general derived based on the Sheth-Tormen mass function introduced above. For the linear halo bias one obtains then

$$b_1(m, z) = 1 + \frac{q[v(m, z)]^2 - 1}{\delta_{\text{sc}}(z)} + \frac{2p}{\delta_{\text{sc}}(z)[1 + (q[v(m, z)]^2)^p]}, \quad (34)$$

where  $p$  and  $q$  match the values used in the mass function. Expressions for higher-order bias factors can be found, e.g., in Scoccimarro et al. (2001). Since they only have a small impact on the quantities employed here, we take into account only the first-order bias. In Fourier space we may then write

$$\tilde{\delta}_h(\tilde{\delta}) \equiv \tilde{\delta}_h(k; m, z) = b_1(m, z)\tilde{\delta}(k). \quad (35)$$

6. To obtain the final correlation function, one has to perform integrations along the halo mass and optionally along the halo concentration, with limits formally extending from 0 to  $\infty$ . In practice, we use the mass limits  $m_{\text{min}} = 10^3 h^{-1} M_\odot$  and  $m_{\text{max}} = 10^{16} h^{-1} M_\odot$ . Masses smaller than  $m_{\text{min}} = 10^3 h^{-1} M_\odot$  give no significant contribution to the considered quantities, while, due to the exponential cut-off in mass, masses larger than  $m_{\text{max}} = 10^{16} h^{-1} M_\odot$  are rare. For the concentration, we employ the integration limits  $c_{\text{min}} = 1$  and  $c_{\text{max}} = 10^3$ .
7. Due to the cut-off in mass, the consistency relation (Scoccimarro et al. 2001)

$$\frac{1}{\bar{\rho}} \int_{m_{\text{min}}}^{m_{\text{max}}} dm m n(m, z) b_1(m, z) = 1 \quad (36)$$

does not hold. To cure this problem we consider a rescaled linear bias such that  $b_1(m, z) \rightarrow b_1(m, z)/b_{\text{norm}}(z)$ , where  $b_{\text{norm}}(z)$  is the result of the integral in Eq. (36). In this way, one ensures that the halo term with the largest contribution to the correlation equals the perturbation theory expression on large scales (see Fig. 1).

#### 4.3. Building Blocks

Using the ingredients described in the previous section, it is possible to define building blocks, which simplify significantly the notation for expressing the polyspectra (Cooray & Hu 2001):

$$M_{ij}(k_1, \dots, k_j; z) \equiv \int_{m_{\text{min}}}^{m_{\text{max}}} dm \int_{c_{\text{min}}}^{c_{\text{max}}} dc n(m, z) p(c|m) \left( \frac{m}{\bar{\rho}} \right)^j b_i(m, z) [\tilde{u}(k_1; m, c) \cdots \tilde{u}(k_j; m, c)]. \quad (37)$$

In the case  $i = 0$ , we additionally define  $b_0 \equiv 1$ , for consistency.

#### 4.4. Power spectrum

We can now compute the power spectrum from Eq. (20). The result consists of two terms, the 1-halo and the 2-halo terms,  $P_\delta(k) = P_{1h}(k) + P_{2h}(k)$  (Seljak 2000). They are given by

$$P_{1h}(k) = \frac{1}{\bar{\rho}^2} \int_{m_{\min}}^{m_{\max}} dm n(m) m^2 \int_{c_{\min}}^{c_{\max}} dc p(c|m) |\tilde{u}(k; m, c)|^2, \quad (38)$$

$$P_{2h}(k) = \left[ \frac{1}{\bar{\rho}} \int_{m_{\min}}^{m_{\max}} dm n(m) m b_1(m) \int_{c_{\min}}^{c_{\max}} dc p(c|m) \tilde{u}(k; m, c) \right]^2 P_{\text{lin}}(k), \quad (39)$$

where  $P_{\text{lin}}(k)$  denotes the linear perturbation theory power spectrum, defined in Eq. (4), and we use the ingredients summarized earlier-on. Note that, for convenience, we omit the redshift-dependence in the notation. Using the building blocks from Eq. (37), these terms can be written in the following compact form :

$$P_{1h}(k) = M_{02}(k, k), \quad P_{2h}(k) = [M_{11}(k)]^2 P_{\text{lin}}(k). \quad (40)$$

The 1-halo term,  $P_{1h}$ , denotes correlations in space between two points in the same halo, whereas the 2-halo term,  $P_{2h}$ , takes into account correlations between two different halos. Hence, the 1-halo term is dominant on small scales and the 2-halo term is dominant on large scales. Note that the 2-halo term converges to the linear power spectrum on large scales because of the consistency relation of the first-order halo bias factor (see Eq. 36) and of the limit  $\tilde{u}(k; m, c) \rightarrow 1$  for  $k \rightarrow 0$ .

#### 4.5. Trispectrum

We compute now the dark matter trispectrum in the halo model approach. As discussed in Sect. 3, only parallelogram configurations of the trispectrum wave-vectors contribute to the covariance of the convergence power spectrum. Restricting our calculations to these configurations, we obtain four different halo term contributions,

$$T_\delta(k_1, k_2, \cos \varphi) = T_{1h} + T_{2h} + T_{3h} + T_{4h}, \quad (41)$$

which are simpler than in the general case. In addition, we neglect terms involving higher-order halo bias factors since, on most scales, they provide only a small correction (Ma & Fry 2000; Takada & Jain 2003). We further note that the perturbative expansion of halo centers, used in the calculations, was shown to become inaccurate on non-linear scales (Smith et al. 2007). The contributions to the trispectrum take the following forms, using the compact notation of the building blocks (see Cooray & Sheth 2002 for the expression of the halo model trispectrum including higher-order bias factors): The 1-halo term, dominant on the smallest scales, is

$$T_{1h} = M_{04}(k_1, -k_1, k_2, -k_2). \quad (42)$$

The 2-halo term has two contributions,  $T_{2h} = T_{2h}^{31} + T_{2h}^{22}$ , consisting of a term  $T_{2h}^{31}$ , which corresponds to correlations of three points within one halo and a fourth point in a second halo, and a term  $T_{2h}^{22}$ , which describes correlations involving two points in a first halo and the other two points in the second halo. They read,

$$T_{2h}^{31} = 2M_{13}(k_1, k_2, k_2)M_{11}(k_1)P_{\text{lin}}(k_1) + 2M_{13}(k_1, k_1, k_2)M_{11}(k_2)P_{\text{lin}}(k_2), \quad (43)$$

$$T_{2h}^{22} = M_{12}^2(k_1, k_2)[P_{\text{lin}}(|\mathbf{k}_1 + \mathbf{k}_2|) + P_{\text{lin}}(|\mathbf{k}_1 - \mathbf{k}_2|)]. \quad (44)$$

The 3-halo term is given by

$$T_{3h} = 2M_{12}(k_1, k_2)M_{11}(k_1)M_{11}(k_2) \left[ B_{\text{pt}}(\mathbf{k}_1, \mathbf{k}_2, -\mathbf{k}_1 - \mathbf{k}_2) + B_{\text{pt}}(\mathbf{k}_1, -\mathbf{k}_2, -\mathbf{k}_1 + \mathbf{k}_2) \right]. \quad (45)$$

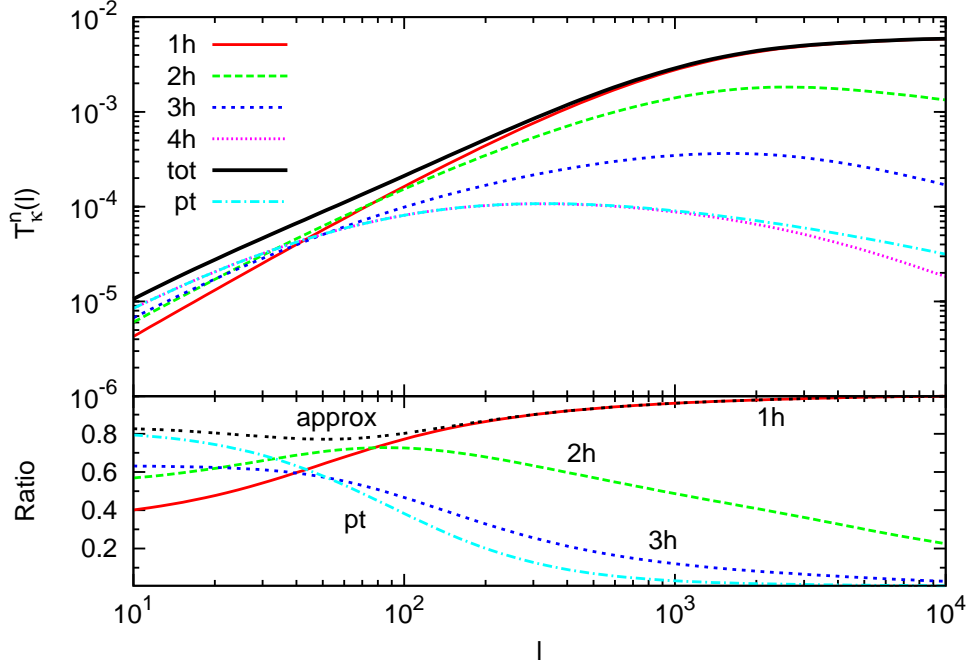
Finally, the 4-halo term, dominant on large scales, is

$$T_{4h} = M_{11}^2(k_1)M_{11}^2(k_2)T_{\text{pt}}(\mathbf{k}_1, -\mathbf{k}_1, \mathbf{k}_2, -\mathbf{k}_2), \quad (46)$$

and describes correlations of points distributed in four different halos. Note that, like for the power spectrum, the 2-halo term is computed from the linear power spectrum. On the other hand, the 3- and 4-halo terms depend on  $B_{\text{pt}}$  and  $T_{\text{pt}}$ , respectively, which are the lowest-order, non-vanishing, perturbation theory contributions to the bispectrum and trispectrum. Both spectra are derived in Appendix A.2.

#### 4.6. Convergence spectra

The convergence power spectrum and trispectrum, needed to evaluate the covariance in Eq. (14), are computed by projecting  $P_\delta$  and  $T_\delta$ , according to Eqs. (10) and (11). Fig. 1 shows the dimensionless convergence trispectrum, defined as  $\mathcal{T}_\kappa^n(\ell) = \ell^2 / 2\pi \sqrt[3]{T_\kappa^n(\ell, -\ell, \ell, -\ell)}$  for a square configuration where all wave-numbers have a length  $\ell$ , where  $n \in \{1h, 2h, 3h, 4h\}$  denotes the contribution from the corresponding halo term. The plot shows the individual contributions to the dimensionless  $\mathcal{T}_\kappa^n$  (the projections of Eqs. 42-46), illustrating on which scales the individual terms are important, as well as the total contribution of all halo terms  $\mathcal{T}_\kappa^{\text{tot}}$ . We show, in addition, the dimensionless projected  $\mathcal{T}_\kappa^{\text{pt}}$ , which closely follows the 4-halo term. We see that the commonly used approximation  $\mathcal{T}_\kappa^{\text{tot}} \approx \mathcal{T}_\kappa^{\text{pt}} + \mathcal{T}_\kappa^{1h}$  is accurate for large wave-numbers ( $\ell \gtrsim 10^3$ ) but has a deviation of about 20% from the complete trispectrum  $\mathcal{T}_\kappa^{\text{tot}}$  for small wave-numbers ( $\ell \lesssim 10^2$ ).



**Fig. 1.** Square configuration of the dimensionless convergence trispectrum  $\mathcal{T}_\kappa^n(\ell)$  against wave-number  $\ell$  for the range  $10 \leq \ell \leq 10^4$  for  $n \in \{1h, 2h, 3h, 4h, \text{tot}, \text{pt}\}$ . The upper panel displays the four individual halo terms, the sum of all four trispectrum halo terms  $\mathcal{T}_\kappa^{\text{tot}}$  (solid line) and the tree-level perturbation theory trispectrum  $\mathcal{T}_\kappa^{\text{pt}}$ , as indicated in the key. The lower panel shows the ratio between the indicated contributions and the complete trispectrum. The double dashed line illustrates the corresponding ratio of the approximation  $\mathcal{T}_\kappa^{\text{tot}} \approx \mathcal{T}_\kappa^{\text{pt}} + \mathcal{T}_\kappa^{1h}$ . Note that we consider the 4-halo term only in the upper panel since it resembles the term  $\mathcal{T}_\kappa^{\text{pt}}$  on large scales.

#### 4.7. Stochastic halo concentration

The previous results were computed using the *deterministic* concentration-mass relation of Eq. (30). We now analyze the impact of scatter in the halo concentration parameter  $c$  on the covariance of the convergence power spectrum, using the *stochastic* concentration relation given by the log-normal concentration distribution of Eq. (31).

Cooray & Hu (2001), analyzed the effect of a stochastic concentration on the three-dimensional power spectrum and trispectrum and found that the behavior of the corresponding 1-halo terms were increasingly sensitive to the width of the concentration distribution for smaller wave-numbers  $k$ . Furthermore, the effect of a stochastic concentration relation was more pronounced for the trispectrum than for the power spectrum, since the tail of the concentration distribution is weighted more strongly in higher-order statistics.

Performing the same analysis for the projected power spectrum and trispectrum, we find a similar trend as in the three-dimensional case, but with a smaller sensitivity to the concentration width of the distribution on small scales. For  $\sigma_{\ln c} = 0.3$ , we find, in the case of the 1-halo term of the power spectrum, a deviation from a deterministic concentration relation of about 1% – 2% for wave-numbers larger than  $\ell \sim 10^3$ , whereas, for the 1-halo term of the trispectrum, the deviation is of the order of 10% – 15% in the same  $\ell$ -range. Thus, when considering the covariance of the convergence power spectrum, one should take into account the concentration dispersion in the 1-halo term of the trispectrum but can safely neglect it for the power spectrum. Additionally, we find that a stochastic concentration has only a small impact on the 2-halo terms of the power spectrum and trispectrum (Pielorz 2008).

From this analysis, we expect the effect of a concentration distribution to be the strongest on the non-Gaussian part of the covariance, which depends on the trispectrum. To directly infer the impact of a concentration distribution on the covariance, we calculate the 1-halo contribution to the non-Gaussian covariance, i.e., we perform the bin averaging of Eq. (15),

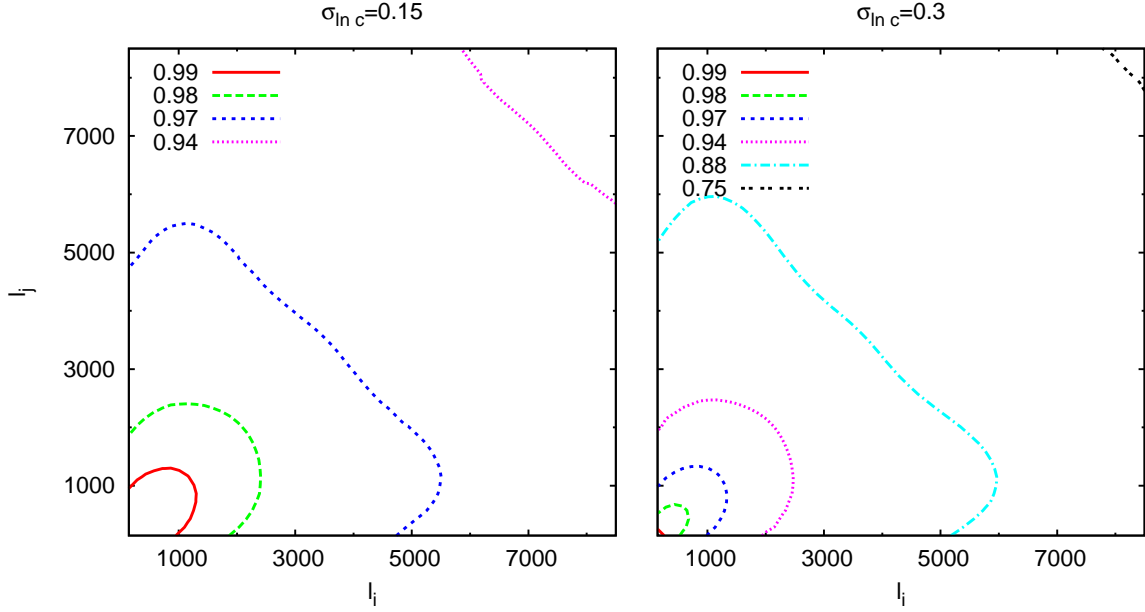
$$C_{1h}^{\text{NG}}(\ell_i, \ell_j) \simeq \int_{|\ell_1| \in \ell_i} \frac{d\ell_1}{A_r(\ell_i)} \ell_1^3 \int_{|\ell_2| \in \ell_j} \frac{d\ell_2}{A_r(\ell_j)} \ell_2^3 T_\kappa^{1h}(\ell_1, \ell_2), \quad (47)$$

for two different concentration dispersions  $\sigma_{\ln c}$ , using our halo model implementation. Fig. 2 shows the ratio

$$R(\ell_i, \ell_j) \equiv \frac{C_{1h}^{\text{NG}}(\ell_i, \ell_j; \sigma_{\ln c})}{C_{1h}^{\text{NG}}(\ell_i, \ell_j; \sigma_{\ln c} = 0)}, \quad (48)$$

where  $\sigma_{\ln c} = 0$  denotes the deterministic concentration relation, for two values  $\sigma_{\ln c} = \{0.15, 0.3\}$  in two contour plots. In agreement with the previous results, the largest impact of the concentration dispersion occurs for wave-numbers larger than  $\ell_i \simeq \ell_j \simeq 4000$ . For  $\sigma_{\ln c} = 0.15$ , the deviation of the covariance  $C_{1h}^{\text{NG}}$  from the original deterministic concentration relation is small, of about 3% – 6%. The effect becomes non-negligible for  $\sigma_{\ln c} = 0.3$ , where we find a deviation of 12% (for  $\ell_i \simeq \ell_j \gtrsim 4000$ ) to 25% (for  $\ell_i \simeq \ell_j \gtrsim 8000$ ).





**Fig. 2.** Contour plots of the ratio  $R(\ell_i, \ell_j)$  between the non-Gaussian covariance 1-halo term contribution (see Eq. 48) computed with a concentration dispersion  $\sigma_{\ln c}$  and with a deterministic concentration, against wave-numbers  $(\ell_i, \ell_j)$  ranging from  $\ell_i = \ell_j = 500$  to  $\ell_i = \ell_j = 8500$ . The left panel is for concentration dispersion  $\sigma_{\ln c} = 0.15$ , whereas in the right panel  $\sigma_{\ln c} = 0.3$ .

**Table 1.** Cosmological parameters used to set up the initial power spectrum, which determines how the simulation particles are distributed initially. The simulations employ either the BBKS (Bardeen et al. 1986) or the EH (Eisenstein & Hu 1998) transfer function. Convergence maps with source galaxies situated at  $z_s = 1$  and  $z_s = 2$  were produced for both sets of simulations.

Simulation	$\Omega_m$	$\Omega_\Lambda$	$h$	$\Omega_b$	$\sigma_8$	$n_s$	$\Gamma$	$z_s$	$T(k)$
Virgo	0.3	0.7	0.7	0.0	0.9	1.0	0.21	1 (2)	BBKS
Gems	0.25	0.75	0.7	0.04	0.78	1.0	0.14	1 (2)	EH

**Table 2.** Parameters used for generating the two  $N$ -body simulations considered in this paper and for producing the resulting convergence maps with the multiple-lens-plane ray-tracing algorithm (see e.g., Jain et al. 2000; Hilbert et al. 2008). From left to right these are: the side length,  $L_{\text{box}}$ , of the cubic simulation box, the number of particles,  $N_{\text{par}}$ , used for the simulation, their mass,  $m_{\text{par}}$ , and the number of available convergence maps,  $N_{\text{map}}$ , with area  $A_{\text{map}}$ .

Simulation	$L_{\text{box}} / h^{-1} \text{ Mpc}$	$N_{\text{par}}$	$m_{\text{par}} / h^{-1} M_\odot$	$N_{\text{map}}$	$A_{\text{map}} / (\text{deg})^2$
Virgo	141.3	$256^3$	$1.4 \times 10^{10}$	200	0.25
Gems	150.0	$256^3$	$1.4 \times 10^{10}$	220	16.00

## 5. Comparison with $N$ -body simulations

In Sect. 4, we computed the power spectrum and trispectrum of the dark matter fluctuations. Projecting them according to Eqs. (10) and (11), and inserting the result in Eqs. (12), (14), and (15), we obtained the covariance of the convergence power spectrum estimator.

To test the accuracy of the halo model predictions for the statistics of the dark matter density field, we compare the dimensionless convergence power spectrum and the corresponding covariance, calculated in the halo model approach, with results from two different ray-tracing simulations.

### 5.1. Virgo and Gems simulation

For our comparison, we chose one simulation from Jenkins et al. (1998) and ten simulations from Hartlap et al. (2009), which we denote in the following as Virgo and Gems simulation, respectively. The Virgo simulation was carried out in 1997 by the Virgo-Consortium for a  $\Lambda$ CDM cosmology (see Tab. 1) with  $N_{\text{par}} = 256^3$  particles in a periodic box of comoving side length  $L_{\text{box}} = 141.3 h^{-1} \text{ Mpc}$  (see Tab. 2). It uses the PP/PM-code HYDRA, which places subgrids of higher resolution in strongly clustered regions. Structures on scales larger than  $2l_{\text{soft}} \approx 40 h^{-1} \text{ kpc}$  can be considered as well resolved. The Gems simulations

were set up in cubic volumes of comoving side length  $L_{\text{box}} = 150 h^{-1} \text{ Mpc}$  with  $256^3$  particles (see Tab. 2). The cosmology chosen reflects the WMAP5 results (Komatsu et al. 2009) and thus has a lower value for  $\sigma_8$  than the Virgo simulation. It uses the GADGET-2 code to simulate the evolution of dark matter particles (Springel 2005) and has a softening length of  $2l_{\text{soft}} \approx 30 h^{-1} \text{ kpc}$ .

### 5.2. Ray-tracing

The output of numerical simulations are three-dimensional distributions of  $N_{\text{par}}$  particles in cubic boxes over a range of redshift values. In order to compare the results with the predicted convergence power spectrum from the halo model, we make use of the *multiple-lens-plane ray-tracing* algorithm (see e.g., Jain et al. 2000; Hilbert et al. 2008) to construct effective convergence maps. The basic idea is to introduce a series of lens planes perpendicular to the central line-of-sight of the observer's backward light cone. In this way, the matter distribution within the light cone is sliced and can be projected onto the corresponding lens plane. By computing the deflection of light rays and its derivatives at each lens plane, one simulates the photon trajectory from the observer to the source by keeping track of the distortions of ray bundles. In this way, the continuous deflection of light rays is approximated by a finite number of deflections at the lens planes. As a result, one obtains the Jacobian matrix for the lens mapping from source to observer and can construct convergence maps.

For both simulations, a similar number of around 200 effective convergence maps were produced, with source galaxies situated at a single redshift of either  $z_s = 1$  or  $z_s = 2$  (see Tab. 2). The maps produced with the Gems simulation have an area of  $16 \text{ deg}^2$ , while the ones from Virgo are much smaller, with  $0.25 \text{ deg}^2$ .

### 5.3. Convergence power spectrum

To test the accuracy of the halo model approach in describing the non-linear evolution of dark matter, we compare the dimensionless projected power spectrum, computed in the halo model approach, to the ones estimated from the numerical  $N$ -body simulations.

The dimensionless convergence power spectra of the simulations are measured from the real-space two-dimensional convergence maps of length  $L_{\text{map}}$  and grid-size  $N_{\text{bin}}$ . For this, we first apply a Fast Fourier Transform<sup>2</sup> to obtain  $\tilde{\kappa}(\ell)$  on each grid-point. Then, we estimate the power spectrum at a wave-number  $\ell$  for the  $k$ -th convergence map by averaging over all Fourier modes in the band  $\ell_b = \{\ell \mid \ell - \Delta\ell/2 \leq |\ell| \leq \ell + \Delta\ell/2\}$ , i.e.,

$$\hat{\mathcal{P}}^{(k)}(\ell) = \frac{1}{N_p(\ell)} \sum_{\ell \in \ell_b} \frac{\ell^2}{2\pi} |\tilde{\kappa}(\ell)|^2, \quad (49)$$

where  $N_p(\ell)$  denotes the number of modes in each band  $\ell_b$  of width  $\Delta\ell$ . Finally, the ensemble average of the dimensionless power spectrum is obtained by averaging over the results of the various convergence maps, i.e.,  $\mathcal{P}(\ell) = \frac{1}{N_{\text{map}}} \sum_{k=1}^{N_{\text{map}}} \hat{\mathcal{P}}^{(k)}(\ell_i)$ . The error bars for the power spectrum estimate are computed from the dispersion over the  $N_{\text{map}}$  convergence maps used, and are due to sample variance. Since modes with a length similar to the side length of the convergence map are only poorly represented, the sampling variance is largest for small  $\ell$ . The effect is stronger for the Virgo simulation than for the Gems simulation, since the Gems convergence maps have a much larger area  $A_{\text{map}}$ .

Fig. 3 shows a good agreement between the halo model and the  $N$ -body simulation convergence power spectra. We also include, in Fig. 3, the convergence power spectra computed using the fitting formulae of Peacock & Dodds (1996) and Smith et al. (2003). Both, the halo model and the fitting formulae, show a better agreement with simulations for the lower source redshift case. On small scales, the halo model and the Smith et al. fitting formula agree well with the simulations, whereas the Peacock-Dodds fitting formula has too little power, in particular in the case of the Gems simulation. On intermediate scales ( $\ell \approx 10^3$ ), the halo model is less accurate than the Smith et al. fitting formula, suggesting that the halo model suffers from the halo exclusion problem on these scales, as described, e.g., in Tinker et al. (2005). This means that, while in simulations halos are never separated by distances smaller than the sum of their virial radii, this is not accounted for in the framework of our halo model and is probably the cause for the deviation. The good agreement of the Smith et al. formula is not surprising, since it is based on simulations of similar resolution than the ones we consider here. In contrast, a similar comparison using the convergence power spectrum estimated with the Millennium Run (Hilbert et al. 2008) clearly favors the halo model prediction over of the two fitting formulae, with both fitting formulae strongly underestimating the power on intermediate and small scales.

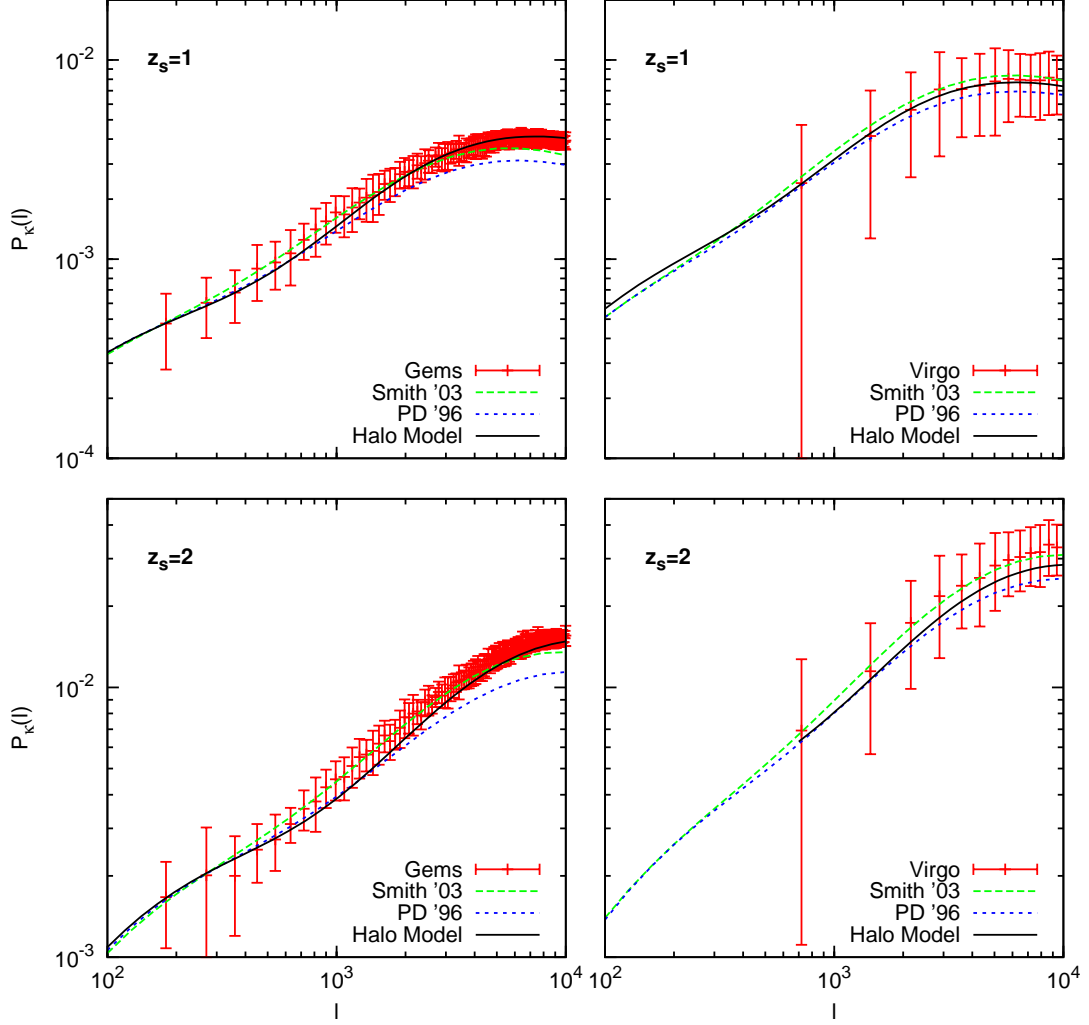
The good overall accuracy of the halo model results was expected since its ingredients, such as the mass function and the halo profile, were obtained from  $N$ -body simulations. We note that, in this analysis, we used a deterministic concentration parameter, since the use of a stochastic one has only a small effect on the small scales of the convergence power spectrum.

### 5.4. Covariance of the convergence power spectrum

The similarity between simulation and halo model power spectra was to some extent expected. The ability to make an accurate description of higher-order correlations provides a stronger test of the halo model. Due to its important role in calculating the error of the power spectrum and for parameter estimates, we focus in this section on the accuracy of the covariance of the dimensionless convergence power spectrum.

We need an appropriate estimator for the power spectrum covariance of the simulations. As each simulation provides  $N_{\text{map}}$  different  $\kappa$ -maps, we have  $N_{\text{map}}$  realizations of the power spectrum. From these we estimate the covariance by applying the unbiased

<sup>2</sup> For the FFT we use an algorithm from the GNU scientific library (see <http://www.gnu.org/software/gsl/> for more details).



**Fig. 3.** Dimensionless convergence power spectrum  $\mathcal{P}_\kappa(\ell)$  against wave-number  $\ell$  for galaxy sources at redshift  $z_s = 1$  (upper panels) and  $z_s = 2$  (lower panels). Points with error bars show measurements from the two sets of numerical simulations: Gems (left plots) and Virgo (right plots). Both simulations use a similar particle setup but differ in the area of the  $\kappa$ -maps (see Tabs. 1 and 2). They are compared with the corresponding results obtained with fitting formulae from Smith et al. (long-dashed line), Peacock-Dodds (short-dashed line), and the halo model predictions for a deterministic halo concentration (solid line). See text for a discussion.

sample covariance estimator which has for our purpose the form:

$$C_{\text{sim}}(\ell_i, \ell_j) = \frac{1}{N_{\text{map}} - 1} \left[ \sum_{k=1}^{N_{\text{map}}} \hat{\mathcal{P}}_\kappa^{(k)}(\ell_i) \hat{\mathcal{P}}_\kappa^{(k)}(\ell_j) - \frac{1}{N_{\text{map}}} \left( \sum_{k=1}^{N_{\text{map}}} \hat{\mathcal{P}}_\kappa^{(k)}(\ell_i) \right) \left( \sum_{k=1}^{N_{\text{map}}} \hat{\mathcal{P}}_\kappa^{(k)}(\ell_j) \right) \right], \quad (50)$$

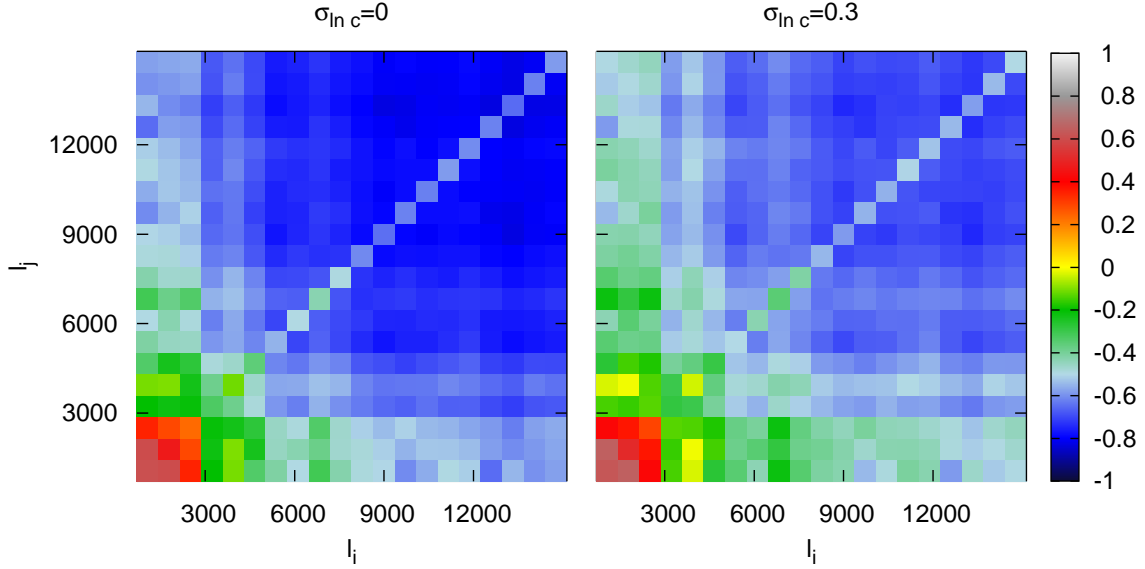
where  $\hat{\mathcal{P}}_\kappa^{(k)}(\ell_i)$  is the projected power spectrum estimate of the  $k$ -th effective convergence map at a wave-number  $\ell_i$  (see Eq. 49). We evaluate  $C_{\text{sim}}$  for both, Virgo and Gems simulations, for the case  $z_s = 1$ .

The halo model results,  $C_{\text{halo}}$ , are calculated as described in Sect. 4 and include all terms of the non-Gaussian covariance. In their computation, we use our fiducial halo model with the ingredients summarized in Sect. 4.2, and use the cosmological parameters values corresponding to each simulation, given in Tab. 1, for the case  $z_s = 1$ . We consider both, a deterministic concentration-mass relation and a stochastic one, with dispersion  $\sigma_{\text{inc}} = 0.3$  for the 1-halo term of the trispectrum.

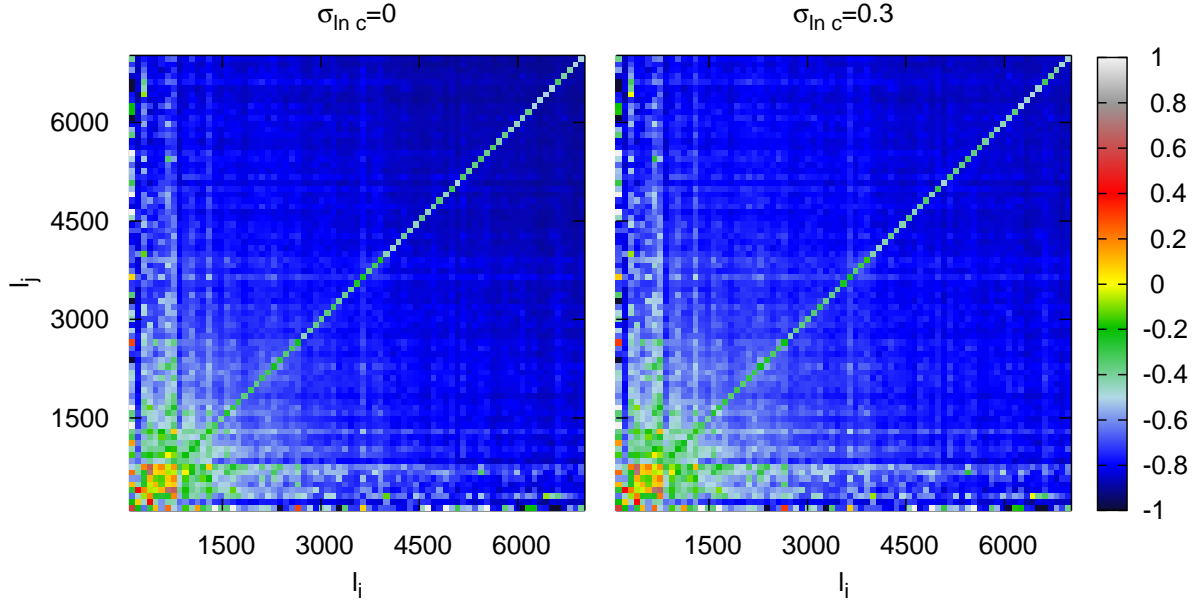
We compare the halo model with the simulations covariance matrices considering their relative deviation

$$\left( \frac{\Delta C_{ij}}{C_{ij}} \right) = \frac{C_{\text{halo}}(\ell_i, \ell_j) - C_{\text{sim}}(\ell_i, \ell_j)}{C_{\text{sim}}(\ell_i, \ell_j)}. \quad (51)$$

We note that although the number of available convergence maps for our simulations ( $\sim 200$ ) is too small to obtain a percentage level accuracy for the covariance estimate (Takahashi et al. 2009), we assume that the covariance from the simulations is the reference

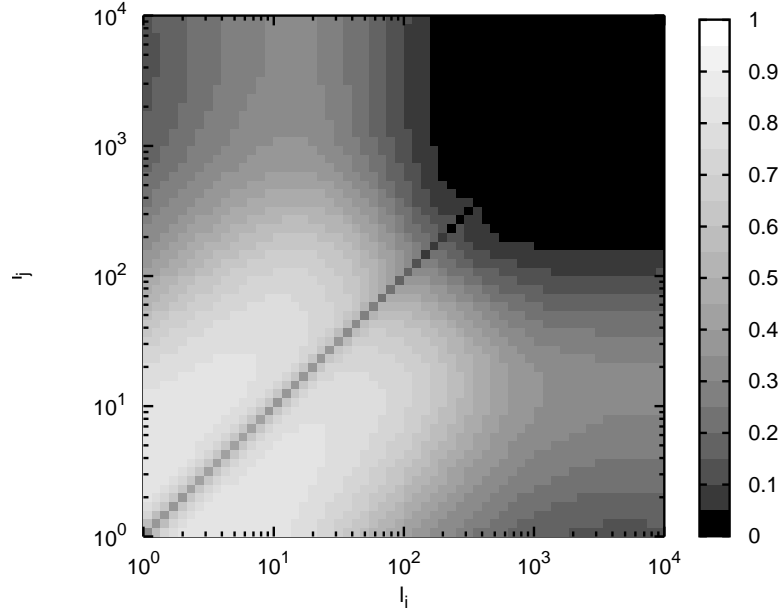


**Fig. 4.** Relative difference  $\Delta C_{ij}/C_{ij}$  (see Eq. 51) between the halo model prediction for the covariance of the convergence power spectrum and the results from the Virgo simulation ( $z_s = 1$ ) against wave-numbers ( $\ell_i, \ell_j$ ). The binning scheme is linear with a constant bin width of  $\Delta\ell = 720$  ranging from  $\ell_0 = 720$  to  $\ell_{20} = 15120$ . The left panel displays the full covariance, including all four halo terms for the trispectrum, and uses a deterministic concentration-mass relation denoted by  $\sigma_{\ln c} = 0$ . The right panel illustrates the same covariance but with a stochastic concentration distribution of width  $\sigma_{\ln c} = 0.3$  for the 1-halo term of the trispectrum.



**Fig. 5.** Relative difference  $\Delta C_{ij}/C_{ij}$  (see Eq. 51) between the halo model prediction for the covariance of the convergence power spectrum and the results from the Gems simulation ( $z_s = 1$ ) against wave-numbers ( $\ell_i, \ell_j$ ). The binning scheme is linear with a constant bin width of  $\Delta\ell = 90$  ranging from  $\ell_0 = 90$  to  $\ell_{20} = 7020$ . The left panel displays the full covariance, including all four halo terms for the trispectrum, and uses a deterministic concentration-mass relation denoted by  $\sigma_{\ln c} = 0$ . The right panel illustrates the same covariance but with a stochastic concentration distribution of width  $\sigma_{\ln c} = 0.3$  for the 1-halo term of the trispectrum.

one, and put it in the denominator of Eq. (51). The comparisons are displayed in Figs. 4 and 5 for the Virgo and Gems simulation, respectively. In the case of the Virgo simulation with  $\sigma_{\ln c} = 0$ , the best agreement is a relative deviation of 20% – 40% found on intermediate scales between  $\ell \simeq 1500$  and  $\ell \simeq 5000$ . For small scales ( $\ell \gtrsim 6000$ ) the halo model underestimates the simulation by 60% or more. On large scales, i.e., for wave-numbers  $\ell \lesssim 1500$ , the halo model overestimates the simulation by around 60%. However, since the simulation covariances suffer from a large sampling variance due to the small size of the convergence maps, the comparison is not meaningful on these scales. The agreement improves for  $\sigma_{\ln c} = 0.3$ , in this case there is a larger range of scales where the deviation is small. For the Gems simulation the agreement is much better. There is now an interval of scales ( $\ell \lesssim 1500$ ) where the deviation is in the range 0% – 40%. Throughout all the scales probed, the deviation on the diagonal of the covariances is



**Fig. 6.** Ratio  $C_{\text{pt}}/C^{\text{NG}}$  against wave-numbers  $(\ell_i, \ell_j)$ . The covariance in tree-level perturbation theory resembles the non-Gaussian covariance on very large scales (i.e.,  $\ell_i = \ell_j \leq 20$ ) to approximately 80% – 90% except along the diagonal and the closest off-diagonals. Here a good approximation of the non-Gaussian covariance requires at least one additional halo term. For wave-numbers  $\ell_i = \ell_j \geq 100$ , an accurate description of the non-Gaussian covariance requires at least the 1-halo term.

**Table 3.** The fiducial model used for the fitting procedure. See text for details.

$\Omega_m$	$\Omega_\Lambda$	$h$	$\Omega_b$	$\sigma_8$	$n_s$	$\sigma_{\text{inc}}$	$T(k)$
0.28	0.72	0.73	0.045	0.82	1.0	0.3	EH

around 20% – 40%. However, for off-diagonal components, at small scales, the agreement is still poor. Finally, the improvement of considering  $\sigma_{\text{inc}} = 0.3$  is less significant than in the Virgo case.

Although the halo model predictions strongly deviate from the simulation estimates of the covariance on small scales, this analysis does not necessarily imply a poor accuracy of those predictions. Indeed, the simulation covariances estimated in this analysis have a strong scatter. In particular, in the case of the Gems simulation, we found from bootstrap subsamples of 50 convergence maps that the resulting covariances can deviate up to 20% from the average covariance of the complete sample (Pielorz 2008). This is in agreement with the results by Takahashi et al. (2009) who need to use 5000 simulations to obtain an estimate of the matter power spectrum covariance at a sub-percent level accuracy.

There are however very recent indications that Eq. (14) indeed underestimates the covariance of the convergence power spectrum on small scales (Sato et al. 2009). In particular, sample variance in the number of halos in a finite field is not accounted for by Eq. (14). Indeed, the mass function yields a mean number density, but there are fluctuations on the number of halos due to the large-scale mass fluctuations. Sample variance in the number of clusters in a volume-limited survey (Hu & Kravtsov 2003) has been considered in cluster abundance studies (e.g., Vikhlinin et al. 2009). In the halo model framework, the sample variance in the number of halos was derived in Takada & Bridle (2007) and its contribution to the covariance of the convergence power spectrum was found, in Sato et al. (2009), to boost the non-Gaussian errors of a 25 deg<sup>2</sup> survey by one order of magnitude on scales  $\ell \approx 10^4$ . The increase is reduced for larger survey areas.

## 6. Fitting formula for the covariance of the convergence power spectrum

Future weak lensing surveys will provide much more precise measurements of the convergence power spectrum. In order to obtain robust constraints on cosmological parameters, accurate estimates of both the power spectrum and its covariance are needed.

### 6.1. Methodology

The number of measured power spectra is, in general, insufficient to infer the complete covariance directly from observations. One has thus to derive it either from ray-tracing maps of numerical  $N$ -body simulations or with an analytic approach. A drawback of the first method is that it requires a large number of realizations and, in addition, is very time-consuming if an exploration of the covariance in the parameter space is needed.

In the previous sections, we derived the covariance, and in particular its non-Gaussian part, with an analytic approach. This computation is, however, time-consuming and it might be useful to obtain an accurate covariance in a faster way. A first approach would be to rely on stronger approximations. For example, a commonly used approximation consists on evaluating the non-Gaussian

**Table 4.** Best-fit values for the parameters  $(p_0, p_1, p_2)$  of the redshift-fit for the set of coefficients  $(a_0, \dots, a_8)$ . The 3-parameter fit is defined in Eq. (55) and is valid in the range  $z_s \in [0.5, 2]$ .

	$p_0$	$p_1$	$p_2$
$a_0$	$4.974 \times 10^{-2}$	0.8385	$-7.877 \times 10^{-3}$
$a_1$	$1.854 \times 10^{-3}$	2.614	$5.926 \times 10^{-3}$
$a_2$	$-2.037 \times 10^5$	2.350	$7.962 \times 10^3$
$a_3$	$-5.564 \times 10^4$	3.152	$-1.625 \times 10^4$
$a_4$	$3.149 \times 10^{-6}$	0.1108	$-2.172 \times 10^{-6}$
$a_5$	$5.349 \times 10^{-8}$	1.332	$-9.222 \times 10^{-8}$
$a_6$	$-2.406 \times 10^{-3}$	3.116	$-6.171 \times 10^{-4}$
$a_7$	$-1.444 \times 10^{-4}$	5.068	$9.201 \times 10^{-5}$
$a_8$	$-4.285 \times 10^{-3}$	2.529	$2.150 \times 10^{-4}$

covariance from  $T_\kappa \approx T_{\text{pt}} + T_{\text{lh}}$ , instead of using the full trispectrum. We saw in Sect. 4 that this approximation recovers the full trispectrum for scales  $\ell \gtrsim 10^3$  but deviates by  $\sim 40\%$  on scales  $\ell \lesssim 10^2$  for square configurations (compare also with Fig. 6).

An alternative approach that we consider in the following, is to find a fitting formula for the halo model covariance that can be subsequently used without the need for implementing the halo model. We will provide a fit only for the non-Gaussian part of the halo model covariance, since the Gaussian part only depends on the non-linear convergence power spectrum and can thus be accurately computed without relying on the halo model. The inclusion of non-Gaussian errors increases the total covariance and one might think of fitting the ratio between the non-Gaussian and Gaussian terms. However, this is not a good quantity to fit, since the Gaussian contribution is diagonal and binning-dependent. In contrast, a similar ratio was fitted in the real space, where the Gaussian term is non-diagonal and binning-independent, using a non-Gaussian contribution measured from  $N$ -body simulations (Semboloni et al. 2007). In Fourier space, there is some relevant analogous information contained in the tree-level perturbation theory trispectrum. Indeed, pursuing the analogy with the real-space fit,  $T_{\text{pt}}$  is a non-diagonal and binning-independent quantity, with a lower amplitude than the full trispectrum, approaching it at large scales. We thus compute  $C_{\text{pt}}$ , the covariance predicted in tree-level perturbation theory on large scales (see Eq. A.27 and Appendix A for a detailed derivation) and compare it with our covariance computed with the full halo model trispectrum,  $C^{\text{NG}}$ , defined in Eq. (14). The ratio between the two covariances is shown in Fig. 6. In agreement with Fig. 1, the covariance predicted by tree-level perturbation theory contributes only  $\sim 50\%$  to the complete non-Gaussian covariance along the diagonal. On very large scales ( $\ell < 100$ ), the ratio  $C^{\text{NG}}/C_{\text{pt}}$  lies between 1.1 and 2. On smaller scales,  $C_{\text{pt}}$  decreases fast and is no longer useful for fitting purposes.

This discussion motivates us to use two different fitting formulae to model the non-Gaussian covariance over the whole range of scales. On large scales ( $1 < \ell < 200$ ), we model the ratio  $C^{\text{NG}}/C_{\text{pt}}$  as a polynomial in the wave-numbers,  $\ell_i, \ell_j$ , and the dimensionless non-linear convergence power spectrum  $\mathcal{P}_\kappa(\ell)$ . More precisely, we assume

$$C_{<}(\ell_i, \ell_j) = C_{\text{pt}}(\ell_i, \ell_j) [1.1 + a_0 \ell_{\min} + a_1 \ell_{\max} + a_2 \mathcal{P}_\kappa(\ell_{\min}) + a_3 \mathcal{P}_\kappa(\ell_{\max})], \quad (52)$$

where  $\ell_{\min} = \min(\ell_i, \ell_j)$  and  $\ell_{\max} = \max(\ell_i, \ell_j)$  which ensures the symmetry property  $C_{<}(\ell_i, \ell_j) = C_{<}(\ell_j, \ell_i)$ . In this way,  $C_{<} \rightarrow 1.1 C_{\text{pt}}$  for large scales as predicted by the halo model (see Fig. 6), whereas the non-linear clustering on smaller scales is encoded in the first-order polynomial in wave-numbers and power spectra. On small scales, i.e. for  $200 < \ell < 8000$ , we model directly the non-Gaussian covariance amplitude using a second-order polynomial in the power spectrum, such that

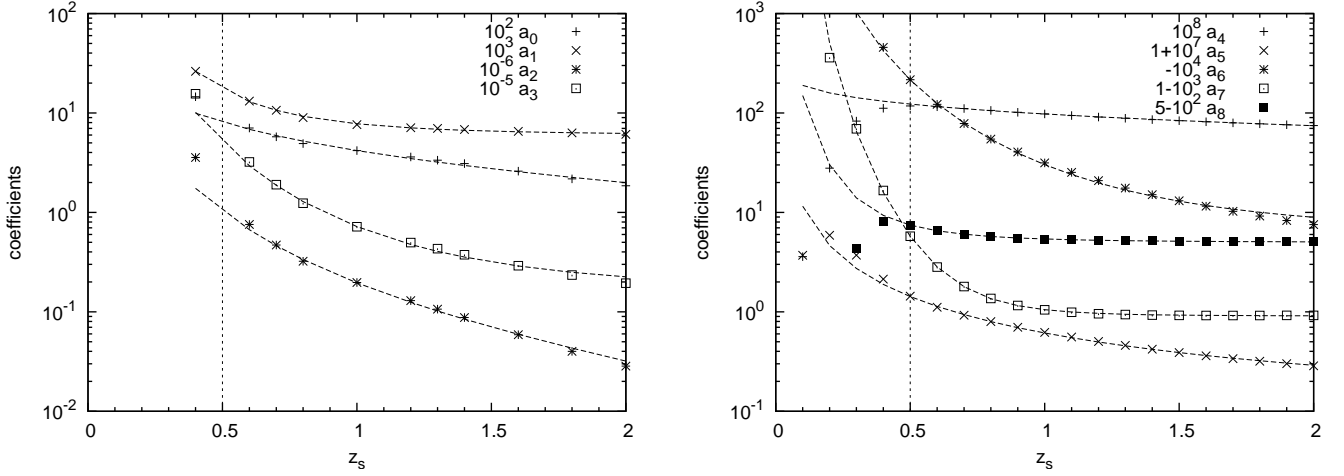
$$C_{>}(\ell_i, \ell_j) = a_4 \mathcal{P}_\kappa(\ell_{\min}) + a_5 \mathcal{P}_\kappa(\ell_{\max}) + a_6 \mathcal{P}_\kappa^2(\ell_{\min}) + a_7 \mathcal{P}_\kappa^2(\ell_{\max}) + a_8 \mathcal{P}_\kappa(\ell_{\min}) \mathcal{P}_\kappa(\ell_{\max}). \quad (53)$$

To ensure a smooth transition from small to large wave-numbers, we consider a linear combination of the two matrices defined in Eqs. (52) and (53) with weightings of third-order in the wave-numbers, such that the full non-Gaussian covariance becomes

$$C_{\text{fit}}^{\text{NG}}(\ell_i, \ell_j) = \frac{\ell_0^3}{\ell_0^3 + (\ell_i + \ell_j)^3} C_{<}(\ell_i, \ell_j) + \frac{(\ell_i + \ell_j)^3}{\ell_0^3 + (\ell_i + \ell_j)^3} C_{>}(\ell_i, \ell_j), \quad (54)$$

with the transition scale fixed at  $\ell_0 = 600$ . We use thus a model with nine free parameters, which is a compromise between expressive power and simplicity, and perform the fit using the range  $1 \leq \ell \leq 8000$ . More precisely, we find the coefficients for the final fit formula by performing two least square fits to Eq. (52) and Eq. (53), respectively. To obtain the coefficients for small (large) scales we computed the corresponding covariance as described above within the halo model approach with 61 bins in the range  $10^2 \leq \ell \leq 10^4$  ( $1 \leq \ell \leq 10^3$ ) and employed the condition  $\ell_i + \ell_j \geq 200$  ( $\ell_i + \ell_j < 200$ ) for the fit.

The covariances used in the fitting procedure, both tree-level perturbation theory and halo model covariance, depend on perturbation theory polyspectra. These were evaluated from the expressions derived in Appendix A using a linear matter power spectrum computed with the Eisenstein-Hu transfer function (Eisenstein & Hu 1998) and assuming the WMAP5-like fiducial model shown in Tab. 3. The non-linear convergence power spectrum, used in the polynomial expressions, was evaluated from the same linear power spectrum. In addition, the halo model non-Gaussian covariance was evaluated using the input parameters as described in Sect. 4.2, including a stochastic concentration-mass relation with  $\sigma_{\ln c} = 0.3$  for the 1-halo term of the trispectrum.



**Fig. 7.** Coefficients  $a_i$  of the fitting formula obtained for various source redshifts. Left (right) panel shows the coefficients of the large (small) scales fitting formula. For convenience, scaled versions of the coefficients are shown, as indicated in the key. Note for example that  $a_5$ ,  $a_7$  and  $a_8$  change sign. The lines show the redshift fit, Eq. (55), with the parameters given in Tab. 4. They provide a good fit for  $z_s > 0.5$  (indicated by the vertical line).

### 6.2. Redshift dependence

All quantities,  $C^{\text{NG}}$ ,  $C_{\text{pt}}$ , and  $\mathcal{P}_\kappa$ , were evaluated for several values of source redshifts (assuming a single source redshift plane), in the range  $0.1 \leq z_s \leq 2.0$ . For each redshift, we perform the fit and find a set of best-fit values for the coefficients ( $a_0, \dots, a_8$ ). Next, we fit the best-fit values of each coefficient as a function of redshift. Some of the best-fit values are increasing functions of the source redshift, while others decrease with redshift, and others still are non-monotonic. They all are, however, monotonic in the redshift range of interest for current and future weak lensing surveys,  $0.5 \leq z_s \leq 2.0$ . In this range, the best-fit values  $a_i(z_s)$  are accurately fitted with

$$a(z_s) = \frac{p_0}{(z_s)^{p_1}} + p_2. \quad (55)$$

Tab. 4 shows the resulting values of the 27 parameters, which define our fitting formula for the halo model covariance of the convergence power spectrum. The behavior of the coefficients with redshift is shown in Fig. 7. We note that a few of them change sign, with most remaining always positive. In addition, their amplitudes may be quite distinct, since they are applied to quantities of quite different amplitudes, such as wave-number, power spectrum or power spectrum squared. The coefficients of the fitting formula for large scales, Eq. (52), all decrease with redshift in absolute values. This compensates the increase of the power spectrum with redshift, allowing for a decrease of the ratio  $C^{\text{NG}}/C_{\text{pt}}$  with redshift, as expected.

### 6.3. Accuracy of the fitting formula

To test the performance of the fitting, we calculate the relative deviation between the non-Gaussian covariance computed with the fitting formula and the halo model one. The deviation,

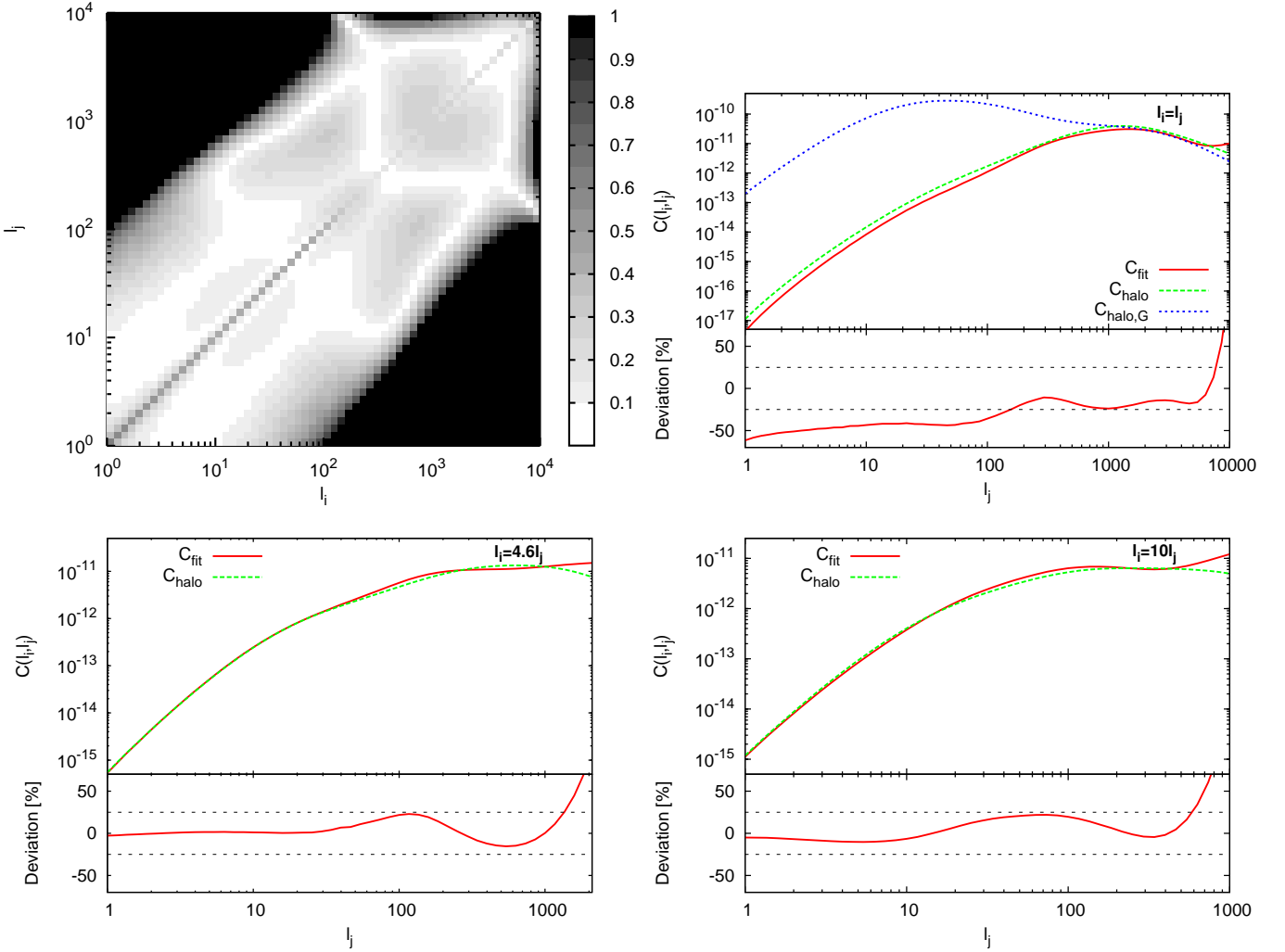
$$\left( \frac{\Delta C_{ij}}{C_{ij}} \right) \equiv \frac{C_{\text{fit}}^{\text{NG}}(\ell_i, \ell_j) - C_{\text{halo}}^{\text{NG}}(\ell_i, \ell_j)}{C_{\text{halo}}^{\text{NG}}(\ell_i, \ell_j)}, \quad (56)$$

is computed for every wave-number pair  $(\ell_i, \ell_j)$  and shown in Fig. 8, for the case of  $z_s = 1$ . The upper left panel of Fig. 8 shows the absolute value of the deviation for each element of the covariance matrix, while the other panels show the deviation along cuts through diagonals of the covariance.

The fit works quite well on the off-diagonal elements that are close to the diagonal, showing an average overestimation of 10%. When moving along any off-diagonal, from larger to smaller scales, (lower panels of Fig. 8), we move from the first fit, Eq. (52), where the deviation is mostly positive, to the second one, Eq. (53), where the deviation is mostly negative. The transition occurs at the local maximum, which indicates that none of the fitting formulae should be extrapolated to the other region. The fits break down on the smallest scales, with the deviation increasing very rapidly when the largest scale reaches  $\ell_i \approx 5000$ .

As we move away from the diagonal the fit gets increasingly worse, in particular the deviations are larger than 100% in the region shown in black in the upper left panel of Fig. 8, which correlates very small with very large scales. The reason for this is that this region was effectively not fitted, since it is not contained in any of the two blocks fitted by Eq. (54). Restricting to the range where both scales are between  $50 < \ell < 5000$ , roughly 90% of the elements show deviations between  $-25\%$  and  $+25\%$ , with the average of the absolute deviations being 10%. This range contains also 9% of outliers where the deviations are larger than  $\pm 25\%$ . The outliers occur in the low-amplitude correlations between the largest ( $50 < \ell < 200$ ) and smallest ( $3000 < \ell < 5000$ ) scales.

The fit is worse on the diagonal than on the first off-diagonals (where  $\ell_i/\ell_j < 50$ ). On the diagonal, the fit always underestimates the covariance. Scales in the range  $50 < \ell < 5000$  show a deviation between  $-40\%$  and  $-10\%$ , with an average of  $-20\%$ . The



**Fig. 8.** Accuracy of the fitting formula for  $z_s = 1$ . Upper left panel: relative deviation  $\Delta C_{ij}/C_{ij}$ , defined in Eq. (56). Other panels: diagonal cuts through the fitted non-Gaussian covariance, showing the covariance on the diagonal  $\ell_i = \ell_j$  (upper-right panel), for  $\ell_i = 4.6\ell_j$  (lower-left) and for  $\ell_i = 10\ell_j$  (lower-right). All three panels show the fitted and the halo model non-Gaussian covariances (as well as the Gaussian for the upper-right panel) as function of the lowest scale  $\ell_j$  in the upper part, and the deviation in percent in the lower part, where the dashed lines mark the 25% level. For the Gaussian covariance we employed a logarithmic binning with 61 bins in a range from  $\ell_{\text{low}} = 1$  to  $\ell_{\text{up}} = 10^5$ .

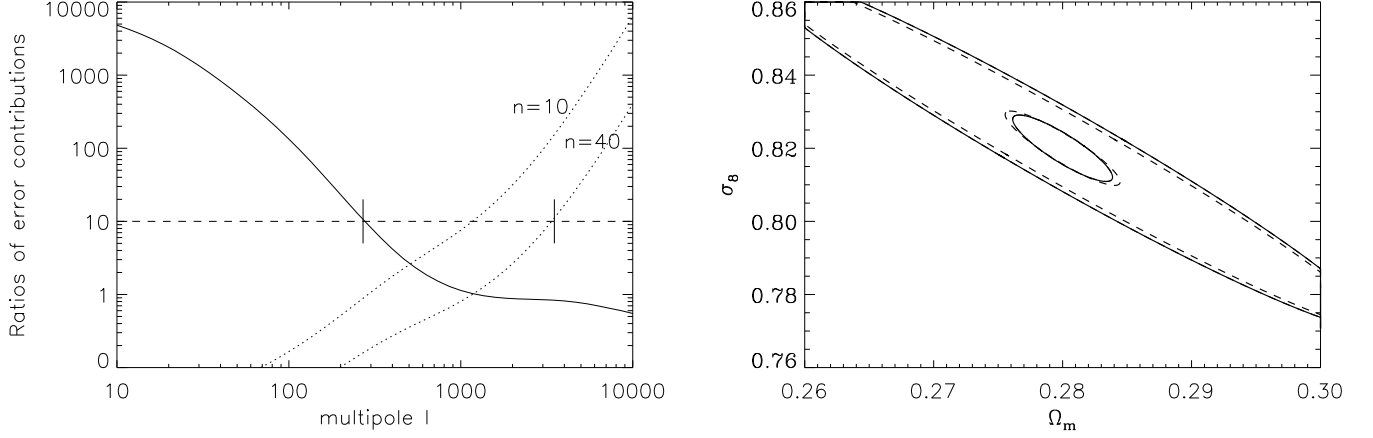
accuracy degrades at larger scales, which is not a problem since for  $\ell < 50$  the non-Gaussian contribution to the total covariance is negligible, as seen in Fig. 8 (upper right panel). Adding the Gaussian contribution to the fitted non-Gaussian one, the underestimation in the diagonal elements is always better than 10%, with an average of 5%.

In summary, the fitting formula for the cosmic variance, including Gaussian and non-Gaussian contributions, has an average accuracy of 10% in the off-diagonal and 5% in the diagonal. It is valid when both scales are in the range  $50 < \ell < 5000$ , corresponding to  $2' < \theta < 5^\circ$  in real space. This is roughly the range used in the latest results from CFHTLS-Wide (Fu et al. 2008). This range includes the scales where non-Gaussianity is relevant, i.e., where the cosmic shear error budget is both dominated by cosmic variance and has important contributions from non-linear clustering (see Sect. 6.4).

#### 6.4. Impact of the accuracy of the fitting formula on parameter estimations

We study the impact of the fitting formula accuracy on the estimation of cosmological parameters, using a Fisher matrix approach. For this, we need to take into account not only the Gaussian and non-Gaussian contributions to the covariance, but also the noise in the observed power spectrum. In practical applications, the convergence field in Eq. (13) is obtained from the observed ellipticities of the source galaxies. The intrinsic ellipticity field (i.e., in the absence of a gravitational lensing effect) is assumed to have zero mean and rms of  $\sigma_e$  per component. This shape noise contaminates the observed power spectrum. Assuming that the intrinsic ellipticities of different galaxies do not correlate, the shape noise contribution to the covariance of the power spectrum is diagonal and given by the new terms arising in Eq. (14) when replacing the power spectrum, in that expression, by the observed one defined as (Kaiser





**Fig. 9.** Left panel: Relative contributions to the diagonal of the convergence power spectrum covariance. The binning-dependent Gaussian to non-Gaussian ratio (solid line; see caption of Fig. 8 for the employed binning scheme) and the shape noise to cosmic variance ratio for two surveys with  $n = 10 \text{ arcmin}^2$  and  $n = 40 \text{ arcmin}^2$  (dotted), are shown as function of multipole  $\ell$ . The ticks at the intersection of the 10% line with the ratio curves enclose the approximate range where the non-Gaussian contribution is important for the EUCLID-like survey. Right panel: Fisher ellipses in the  $(\Omega_m, \sigma_8)$  plane ( $2\sigma$  contours) calculated with the halo model covariance (solid) and the fitted covariance (dashed) for the DES-like (large ellipses) and EUCLID-like (small ellipses) surveys. Compared to the halo covariance case, the fitted covariance error ellipses are decreased (enlarged) by 13% (10%) for the DES (EUCLID) case.

1992),

$$P_{\kappa}^{\text{obs}}(\ell) = P_{\kappa}(\ell) + \frac{\sigma_{\epsilon}^2}{n}, \quad (57)$$

where  $n$  is the number density of source galaxies.

We consider the following three surveys: a medium-deep weak lensing survey covering an area of  $170 \text{ deg}^2$  with  $n = 10 \text{ arcmin}^2$ , like the current CFHTLS-Wide, a wider survey of similar depth covering  $5000 \text{ deg}^2$  with  $n = 10 \text{ arcmin}^2$ , like the planned Dark Energy Survey (DES)<sup>3</sup>, and a wide and deep survey with  $20\,000 \text{ deg}^2$  and  $n = 40 \text{ arcmin}^2$ , like the proposed EUCLID<sup>4</sup>. For all three we assume  $\sigma_{\epsilon} = 0.3$  and compute the covariance using the halo model and the developed fit formula in the range  $50 < \ell < 5000$ . Additionally, we add the covariance of a Gaussian contribution using a logarithmic binning with 61 bins in a range from  $\ell_{\text{low}} = 1$  to  $\ell_{\text{up}} = 10^5$ . The wider surveys will measure correlations on scales larger than  $\ell = 50$ , which we do not consider here for comparison purposes. Note also that for very large scales the flat-sky approximation breaks down.

The left panel of Fig. 9 compares the different terms contributing to the diagonal of the covariance, by showing the Gaussian to non-Gaussian ratio and the shape noise to cosmic variance ratio, where by cosmic variance we denote the sum of the Gaussian and non-Gaussian contributions. The ratio between the Gaussian and non-Gaussian terms is independent of the survey and, for our particular choice of binning, non-Gaussianity starts to affect the diagonal around  $\ell = 300$ , where its amplitude is 10% of the Gaussian amplitude, and dominates from  $\ell \approx 1000$  onwards. Shape noise, including both pure shape noise and the coupling with cosmic variance, also becomes important on small scales, but in a survey-dependent way. It is as large as the cosmic variance on  $\ell \approx 200$  ( $\ell \approx 1000$ ) for surveys with 10 (40) galaxies per arcmin squared. The vertical lines in Fig. 9 (left panel) show, for the EUCLID-like survey, the range where the non-Gaussian contribution to the diagonal is non-negligible, i.e., where it accounts for more than 10% of the cosmic variance while having an amplitude of at least 10% of the shape noise. This range is roughly  $300 < \ell < 3000$ , or approximately  $4' < \theta < 50'$ . This is a rough estimate of the minimum range where the fitting formula is required to have a good accuracy.

In addition, the accuracy of the off-diagonal terms is crucial, since the non-Gaussianity is the sole contribution there. To evaluate the required range of validity of the fitting formula, in a way that includes the off-diagonal elements and is independent of bin width, we define the signal-to-noise ratio (Takada & Jain 2009),

$$\left(\frac{S}{N}\right)^2 = \sum_{ij} \mathcal{P}_{\kappa}(\ell_i) C_{ij}^{-1} \mathcal{P}_{\kappa}(\ell_j). \quad (58)$$

For each survey, we compute the signal-to-noise ratio (SNR) using all scales between  $\ell = 50$ , the largest scale where the fitting is valid, to successive values of  $\ell_{\text{max}}$ . The SNR increases with  $\ell_{\text{max}}$ , as more scales are included in Eq. (58). The increasing rate, however, decreases with increasing  $\ell_{\text{max}}$ , tending to zero when additional scales do not carry additional cosmological information. The value of  $\ell_{\text{max}}$  for which this saturation occurs increases with decreasing shape noise, and gives a good indication of the range where the fitting formula is required to have a good accuracy in order to get accurate estimates of cosmological parameters. We find

<sup>3</sup> <http://www.darkenergysurvey.org/>

<sup>4</sup> <http://www.dune-mission.net/>

that, when increasing  $\ell_{\max}$  from  $\ell_{\max} = 5000$  to  $\ell_{\max} = 10000$ , the SNR increases by a factor of 1.03 for the CFHTLS-like survey and by a factor of 1.09 for the EUCLID-like survey. Hence, even for the latter there is no much gain in reaching  $\ell > 5000$ .

We consider now the Fisher information matrix, which to first order, neglecting the cosmology dependence of the covariance matrix, is given by

$$F_{\alpha\beta} = \sum_{ij} \frac{\partial \mathcal{P}_\kappa(\ell_i)}{\partial p_\alpha} C_{ij}^{-1} \frac{\partial \mathcal{P}_\kappa(\ell_j)}{\partial p_\beta}, \quad (59)$$

where the derivatives are taken w.r.t. a set of cosmological parameters  $p_\alpha$ . The Fisher matrix defines the error ellipsoid in parameter space, with  $(F_{\alpha\alpha}^{-1})^{1/2}$  yielding the  $1\sigma$  marginalized error on  $p_\alpha$ , and  $(1/F_{\alpha\alpha})^{1/2}$  giving the  $1\sigma$  error on  $p_\alpha$  assuming the other parameters are perfectly known.

We compute Eq. (59) for both the fit and halo model covariances, for each of the three surveys. We perform the derivatives at the fiducial model of Tab. 3, varying only 2 cosmological parameters ( $\Omega_m, \sigma_8$ ). For each survey, we compare the areas of the two  $2\sigma$  error ellipses (which defines the inverse of the figure-of-merit) thus obtained. For all three cases, there is a good agreement between the two ellipses.

For the two cases with large shape noise, CFHTLS and DES, we find that the fitting formula underestimates the Fisher ellipse, as compared to the halo model covariance. This is expected because with an enhanced diagonal the correlations between bins are weaker, and the result is dominated by the accuracy of the diagonal, where the fitting formula underestimates the covariance, as we saw earlier on. The deviation is however weak, the areas of the ellipses obtained using the fitting formula are 13% smaller than the halo model result, and the deviation is uniformly distributed on the parameter space (see Fig. 9, right panel, large ellipses). This implies that the deviation on the marginalized constraints is much smaller, and we find that the fitting formula underestimates the errors on both  $\Omega_m$  and  $\sigma_8$  by only 1%, for both surveys. This corresponds to a deviation of 0.1% (0.01%) of the parameters values, for CFHTLS (DES). In contrast, for the EUCLID case, where the covariance has larger correlations, the fitted covariance produced an ellipse slightly larger than the halo model one, by about 10% (see Fig. 9, right panel, small ellipses).

### 6.5. Covariance of real-space estimators

For practical purposes it is sometimes more convenient to study real-space correlations rather than correlations in Fourier space. We therefore define an estimator of a general second-order cosmic shear measure which is related to the convergence power spectrum estimator by

$$\hat{\Gamma}(\theta) = \int_0^\infty \frac{d\ell}{2\pi} W(\ell\theta) \hat{\mathcal{P}}_\kappa(\ell), \quad (60)$$

where  $W(x)$  is an arbitrary weight function. A well-known example of this equation are the shear two-point correlation functions  $\xi_+(\theta)$  and  $\xi_-(\theta)$  with weight functions  $W(\ell\theta) = J_0(\ell\theta)$  and  $W(\ell\theta) = J_4(\ell\theta)$ , respectively (see discussion in Joachimi et al. 2008). Using the definition Eq. (60), we find for the relation between the covariance of the real-space estimators to the covariance of the Fourier-space estimators:

$$\text{Cov}[\hat{\Gamma}(\theta), \hat{\Gamma}(\theta')] = \int_0^\infty \frac{d\ell}{2\pi} W(\ell\theta) \int_0^\infty \frac{d\ell'}{2\pi} W(\ell'\theta') \text{Cov}[\hat{\mathcal{P}}_\kappa(\ell), \hat{\mathcal{P}}_\kappa(\ell')], \quad (61)$$

which is related to the covariance of the dimensionless power spectrum used in the fitting formula by  $(2\pi)^2 \text{Cov}[\hat{\mathcal{P}}_\kappa(\ell), \hat{\mathcal{P}}_\kappa(\ell')] = \ell^2 \ell'^2 \text{Cov}[\hat{\mathcal{P}}_\kappa(\ell), \hat{\mathcal{P}}_\kappa(\ell')]$ . Inserting the result of the dimensionless power spectrum covariance given in Eq. (14) yields

$$\text{Cov}[\hat{\Gamma}(\theta), \hat{\Gamma}(\theta')] = \frac{4\pi}{A} \int_0^\infty \frac{d\ell}{\ell^3} \mathcal{P}_\kappa^2(\ell) W(\ell\theta) W(\ell\theta') + \frac{1}{A} \int_0^\infty \frac{d\ell}{\ell} W(\ell\theta) \int_0^\infty \frac{d\ell'}{\ell'} \bar{T}_\kappa(\ell, \ell') W(\ell'\theta'). \quad (62)$$

We find that the Gaussian part of the real-space covariance is independent of the binning scheme and is non-diagonal in contrast to the covariance in Fourier space.

## 7. Conclusions

We present a fitting formula for the halo model prediction of the non-Gaussian contribution to the covariance of the dimensionless power spectrum of the weak lensing convergence. The formula was constructed assuming a  $\Lambda$ CDM cosmology with WMAP5-like cosmological parameters. In particular, it was obtained for  $\Omega_m = 0.28$ ,  $\sigma_8 = 0.82$  and other parameter values as shown in Tab. 3. It is valid for a scale range of  $50 \lesssim \ell \lesssim 5000$ , corresponding to  $2' < \theta < 5^\circ$  in real space and can be used for surveys with galaxy source redshifts  $z_s \in [0.5, 2]$ . In this range, it reproduces the results of a full implementation of the halo model approach, with a scale-averaged accuracy of 10% in the off-diagonal, and 5% in the diagonal elements. The formula also allows us to recover the halo model ( $\Omega_m, \sigma_8$ ) error ellipses within 15%. The range of validity of the formula and its level of accuracy render it applicable to low shape noise scenarios from next generation weak lensing surveys.

To use the formula, shown in Eq. (54), one needs three quantities :

- The non-Gaussian contribution to the covariance of the convergence power spectrum in tree-level perturbation theory  $C_{\text{pt}}$ , shown in Eq. (A.27). This involves the computation of the convergence trispectrum in tree-level perturbation theory, Eq. (A.26), which requires the calculation of the linear power spectrum and of the  $F_2$  and  $F_3$  coupling functions (Eqs. A.10 and A.12).

- The non-linear convergence power spectrum.
- The 9 coefficients of the fit, which are obtained by inserting the 27 values given in Tab. 4 in Eq. (55), for the required redshift.

The Gaussian contribution, calculated from the non-linear convergence power spectrum as given in Eq. (14), may then be added to the result of the formula, to obtain the total covariance. This is the covariance of the estimator without noise, or the cosmic variance.

The work presented in this paper is based on the assumption that the halo model is a powerful approach to probe non-linear clustering. We tried to test this assumption against results from  $N$ -body simulations, but our comparisons were inconclusive. Indeed, such analysis requires ray-tracing simulations with both large number of convergence maps and large convergence map area. Only then it would be possible to minimize the effect of sampling variance in the simulations, which hindered our attempted tests. Such simulations would also allow us to consider error bars for the estimate of the simulation covariance. There are however indications that the halo model approach underestimates the non-Gaussianity of the covariance and there are attempts to include additional contributions (Takada & Jain 2009; Sato et al. 2009).

The reliability of the halo model for higher-order polyspectra also needs to be studied in more detail. Some work in this direction are the analyses of the impact of the triaxiality of the halo profiles (Smith & Watts 2005; Smith et al. 2006), and halo exclusion effects (Tinker et al. 2005). Moreover, the issue of halo substructure (Dolney et al. 2004) and of the effect of a stochastic concentration parameter have to be understood properly. This paper also addresses this last issue. We analyze the impact of a stochastic concentration parameter on the covariance of the convergence power spectrum. We found that the effect can safely be neglected for the Gaussian contribution, with the convergence power spectrum varying only slightly, and at small scales, for concentration scatters of  $\sigma_{\ln c} \simeq 0.2 - 0.3$ . For the non-Gaussian contribution the effect is more pronounced due to the higher sensitivity of the trispectrum to a stochastic concentration relation. In the case of the 1-halo term of the trispectrum, we find it useful to take into account a concentration dispersion of  $\sigma_{\ln c} \gtrsim 0.3$ . The deviation to a deterministic concentration-mass relation is larger than 12% for wave-numbers  $\ell \gtrsim 3000$ .

Although the fitting formula we obtained provides a more thorough estimate for the non-Gaussian contribution to the power spectrum covariance than the earlier approximation of Semboloni et al. (2007) (global accuracy of  $\sim 20\%$  along the diagonal, which becomes less for the off-diagonals), there is still room for improvements. One drawback of our approach is that it requires the computation of the convergence trispectrum in tree-level perturbation theory. Furthermore, it is only applicable to a small range of WMAP5-like cosmologies and is only valid in the interval  $50 \leq \ell \leq 5000$ . A possible way to avoid these problems and extend the accuracy of the fitting formula might be to construct it entirely from its three-dimensional counterpart, the three-dimensional covariance of the matter power spectrum. This has the advantage that perturbations of different length scales are not additionally mixed due to projection effects, which might allow us to cover a wider range of cosmologies. The desired projected covariance could then be obtained by performing an additional integration along the redshift-space. We will address this issue in a future paper.

*Acknowledgements.* The authors thank Christoph Lampert for invaluable discussions, and Jan Hartlap for providing his ray-tracing simulations of the Gems and Virgo simulations. Together with Martin Kilbinger both provided useful comments to the manuscript. JP would like to thank the TRR33. JR is supported by the Deutsche Forschungsgemeinschaft under the project SCHN 342/7-1 within the Priority Programme SPP 1177 ‘Galaxy Evolution’. IT is supported by the Marie Curie Training and Research Network ‘DUEL’.

## Appendix A: Cosmological Perturbation Theory

On large scales, different Fourier modes evolve independently from each other and thus conserve the Gaussian behavior of the density perturbation field  $\tilde{\delta}(\mathbf{k}, a)$ . It is therefore convenient to work in Fourier space and Fourier transform the fields as well as the non-linear fluid equations (consisting of continuity, Euler and Poisson equation) that describe their evolution in an expanding Universe. Contrary to linear perturbation theory, there is a coupling between different Fourier modes mediated by the coupling function  $\alpha(\mathbf{k}_1, \mathbf{k}_2)$  on smaller scales. In this case, the fluid equations for the density contrast and the irrotational peculiar velocity field  $\theta = \nabla \cdot \mathbf{u}$  in Fourier space are given by (e.g., Bernardeau et al. 2002)

$$a\dot{\tilde{\delta}}(\mathbf{k}, a) + \tilde{\theta}(\mathbf{k}, a) = - \int \frac{d^3 k_1}{(2\pi)^3} \int d^3 k_2 \delta_D(\mathbf{k} - \mathbf{k}_1 - \mathbf{k}_2) \alpha(\mathbf{k}_1, \mathbf{k}_2) \tilde{\theta}(\mathbf{k}_1, a) \tilde{\delta}(\mathbf{k}_2, a), \quad (\text{A.1})$$

$$a\dot{\tilde{\theta}}(\mathbf{k}, a) + \frac{3H_0^2 \Omega_m}{2a} \tilde{\delta}(\mathbf{k}, a) = - \int \frac{d^3 k_1}{(2\pi)^3} \int d^3 k_2 \delta_D(\mathbf{k} - \mathbf{k}_1 - \mathbf{k}_2) \beta(\mathbf{k}_1, \mathbf{k}_2) \tilde{\theta}(\mathbf{k}_1, a) \tilde{\theta}(\mathbf{k}_2, a), \quad (\text{A.2})$$

where we introduced the two fundamental mode coupling functions

$$\alpha(\mathbf{k}_1, \mathbf{k}_2) = \frac{(\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{k}_1}{k_1^2}, \quad \beta(\mathbf{k}_1, \mathbf{k}_2) = \frac{|\mathbf{k}_1 + \mathbf{k}_2|^2 (\mathbf{k}_1 \cdot \mathbf{k}_2)}{2k_1^2 k_2^2}. \quad (\text{A.3})$$

For an Einstein-de Sitter (EdS) cosmology it is possible to find a perturbative ansatz that separates the scale- and time dependencies, whereas for a general  $\Lambda$ CDM model it is impossible to find a separable solution to Eqs. (A.1) and (A.2). However, Scoccimarro et al. (1998) showed that it is possible to find a separable solution in any order if one makes an approximation that is valid at percentage level. One indeed finds then the same recursion relation as in the EdS case. The ansatz is then

$$\tilde{\delta}(\mathbf{k}, a) = \sum_{n=1}^{\infty} D^n(a) \tilde{\delta}_n(\mathbf{k}), \quad \tilde{\theta}(\mathbf{k}, a) = -\dot{a} \sum_{n=1}^{\infty} D^n(a) \tilde{\theta}_n(\mathbf{k}). \quad (\text{A.4})$$

Therefore, we find that the whole information on cosmological parameters is encoded in the growth function due to its dependence on the Hubble parameter (see Eq. 5).

### A.1. Coupling Functions

The  $n$ -th order density contrast and the divergence of the peculiar velocity in Eq. (A.4) is given by

$$\tilde{\delta}_n(\mathbf{k}) = \int \frac{d^3 q_1}{(2\pi)^3} \cdots \frac{d^3 q_{n-1}}{(2\pi)^3} \int d^3 q_n \delta_D(\mathbf{k} - \mathbf{q}_{1\dots n}) F_n(\mathbf{q}_1, \dots, \mathbf{q}_n) \tilde{\delta}_1(\mathbf{q}_1) \cdots \tilde{\delta}_1(\mathbf{q}_n), \quad (\text{A.5})$$

$$\tilde{\theta}_n(\mathbf{k}) = \int \frac{d^3 q_1}{(2\pi)^3} \cdots \frac{d^3 q_{n-1}}{(2\pi)^3} \int d^3 q_n \delta_D(\mathbf{k} - \mathbf{q}_{1\dots n}) G_n(\mathbf{q}_1, \dots, \mathbf{q}_n) \tilde{\delta}_1(\mathbf{q}_1) \cdots \tilde{\delta}_1(\mathbf{q}_n), \quad (\text{A.6})$$

and the  $n$ -th order coupling functions  $F_n$  and  $G_n$  are obtained by the following recursion relations (Jain & Bertschinger 1994)

$$F_n(\mathbf{q}_1, \dots, \mathbf{q}_n) = \sum_{m=1}^{n-1} \frac{G_m(\mathbf{q}_1, \dots, \mathbf{q}_m)}{(2n+3)(n-1)} [(2n+1)\alpha(\mathbf{k}_1, \mathbf{k}_2) F_{n-m}(\mathbf{q}_{m+1}, \dots, \mathbf{q}_n) + 2\beta(\mathbf{k}_1, \mathbf{k}_2) G_{n-m}(\mathbf{q}_{m+1}, \dots, \mathbf{q}_n)], \quad (\text{A.7})$$

$$G_n(\mathbf{q}_1, \dots, \mathbf{q}_n) = \sum_{m=1}^{n-1} \frac{G_m(\mathbf{q}_1, \dots, \mathbf{q}_m)}{(2n+3)(n-1)} [3\alpha(\mathbf{k}_1, \mathbf{k}_2) F_{n-m}(\mathbf{q}_{m+1}, \dots, \mathbf{q}_n) + 2n\beta(\mathbf{k}_1, \mathbf{k}_2) G_{n-m}(\mathbf{q}_{m+1}, \dots, \mathbf{q}_n)], \quad (\text{A.8})$$

where  $\mathbf{k}_1 \equiv \mathbf{q}_1 + \dots + \mathbf{q}_m$  and  $\mathbf{k}_2 \equiv \mathbf{q}_{m+1} + \dots + \mathbf{q}_n$ . The initial conditions for these recursion relations are  $F_1 \equiv 1$  and  $G_1 \equiv 1$ . To get the functions  $F_n^{(s)}$  and  $G_n^{(s)}$  that are symmetric in its arguments, one must perform the following symmetrizing procedure

$$F_n^{(s)}(\mathbf{q}_1, \dots, \mathbf{q}_n) = \frac{1}{n!} \sum_{\pi} F_n(\mathbf{q}_{\pi(1)}, \dots, \mathbf{q}_{\pi(n)}), \quad G_n^{(s)}(\mathbf{q}_1, \dots, \mathbf{q}_n) = \frac{1}{n!} \sum_{\pi} G_n(\mathbf{q}_{\pi(1)}, \dots, \mathbf{q}_{\pi(n)}), \quad (\text{A.9})$$

where the sum is taken over all possible permutations  $\pi$  of the set  $\{1, \dots, n\}$ . These equations enable us to calculate the density contrast in the  $n$ -th order of perturbation theory by using the iterative equations for the coupling functions.

The calculation of the second-order coupling functions is straightforward. The result is

$$F_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) = \frac{5}{7} + \frac{2}{7} \frac{(\mathbf{q}_1 \cdot \mathbf{q}_2)^2}{q_1^2 q_2^2} + \frac{1}{2} \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1 q_2} \left( \frac{q_1}{q_2} + \frac{q_2}{q_1} \right), \quad G_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) = \frac{3}{7} + \frac{4}{7} \frac{(\mathbf{q}_1 \cdot \mathbf{q}_2)^2}{q_1^2 q_2^2} + \frac{1}{2} \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1 q_2} \left( \frac{q_1}{q_2} + \frac{q_2}{q_1} \right). \quad (\text{A.10})$$

The third-order coupling function is given by

$$F_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) = \frac{1}{18} \{ 7\alpha(\mathbf{q}_1, \mathbf{q}_{23}) F_2(\mathbf{q}_2, \mathbf{q}_3) + 2\beta(\mathbf{q}_1, \mathbf{q}_{23}) G_2(\mathbf{q}_2, \mathbf{q}_3) + [7\alpha(\mathbf{q}_{12}, \mathbf{q}_3) + 2\beta(\mathbf{q}_{12}, \mathbf{q}_3)] G_2(\mathbf{q}_1, \mathbf{q}_2) \}, \quad (\text{A.11})$$

where  $\mathbf{q}_{ij} \equiv \mathbf{q}_i + \mathbf{q}_j$ . Employing Eq. (A.9), we find the symmetric function

$$\begin{aligned} F_3^{(s)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) = & \frac{7}{54} [\alpha(\mathbf{q}_1, \mathbf{q}_{23}) F_2^{(s)}(\mathbf{q}_2, \mathbf{q}_3) + \alpha(\mathbf{q}_2, \mathbf{q}_{13}) F_2^{(s)}(\mathbf{q}_1, \mathbf{q}_3) + \alpha(\mathbf{q}_3, \mathbf{q}_{12}) F_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2)] \\ & + \frac{4}{54} [\beta(\mathbf{q}_1, \mathbf{q}_{23}) G_2^{(s)}(\mathbf{q}_2, \mathbf{q}_3) + \beta(\mathbf{q}_2, \mathbf{q}_{13}) G_2^{(s)}(\mathbf{q}_1, \mathbf{q}_3) + \beta(\mathbf{q}_3, \mathbf{q}_{12}) G_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2)] \\ & + \frac{7}{54} [\alpha(\mathbf{q}_{12}, \mathbf{q}_3) G_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) + \alpha(\mathbf{q}_{13}, \mathbf{q}_2) G_2^{(s)}(\mathbf{q}_1, \mathbf{q}_3) + \alpha(\mathbf{q}_{23}, \mathbf{q}_1) G_2^{(s)}(\mathbf{q}_2, \mathbf{q}_3)]. \end{aligned} \quad (\text{A.12})$$

From now on the symmetry superscript “(s)” will be omitted because we will only deal with symmetric coupling functions. For the calculation of the trispectrum in the halo model approach as described in Sect. 4.5, one needs perturbation theory. More precisely, we need the subsequent components

$$\begin{aligned} F_3(\mathbf{q}_1, -\mathbf{q}_1, \mathbf{q}_2) = & \frac{7}{54} [\alpha(\mathbf{q}_1, \mathbf{q}_-) F_2(-\mathbf{q}_1, \mathbf{q}_2) + \alpha(-\mathbf{q}_1, \mathbf{q}_+) F_2(\mathbf{q}_1, \mathbf{q}_2)] \\ & + \frac{4}{54} [\beta(\mathbf{q}_1, \mathbf{q}_-) G_2(-\mathbf{q}_1, \mathbf{q}_2) + \beta(-\mathbf{q}_1, \mathbf{q}_+) G_2(\mathbf{q}_1, \mathbf{q}_2)] \\ & + \frac{7}{54} [\alpha(\mathbf{q}_-, \mathbf{q}_1) G_2(-\mathbf{q}_1, \mathbf{q}_2) + \alpha(\mathbf{q}_+, -\mathbf{q}_1) G_2(\mathbf{q}_1, \mathbf{q}_2)], \end{aligned} \quad (\text{A.13})$$

where we have defined the difference vector  $\mathbf{q}_- \equiv \mathbf{q}_2 - \mathbf{q}_1$  and the sum of the vectors  $\mathbf{q}_+ \equiv \mathbf{q}_1 + \mathbf{q}_2$ .

We already mentioned in the previous section that it is possible to find a solution for an arbitrary cosmology if one makes a small approximation. In the literature one can find closed solutions for the second- and third-order coupling functions. The second-order coupling function changes to

$$F_2(\mathbf{q}_1, \mathbf{q}_2) = \frac{1}{2}(1 + \epsilon) + \frac{1}{2} \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1 q_2} \left( \frac{q_1}{q_2} + \frac{q_2}{q_1} \right) + \left( \frac{1}{2} - \frac{\epsilon}{2} \right) \frac{(\mathbf{q}_1 \cdot \mathbf{q}_2)^2}{q_1^2 q_2^2}, \quad (\text{A.14})$$

where  $\epsilon \approx (3/7)\Omega_m^{-2/63}$  for  $\Omega_m \gtrsim 0.1$  (Bernardeau et al. 2002). For our fiducial choice of  $\Omega_m = 0.3$ , we get  $\Omega_m^{-2/63} \approx 1.039$ . Thus, within a few percent correction to the first and last term, the coupling functions are independent of cosmological parameters.

## A.2. Correlation functions

### A.2.1. Bispectrum

The dark matter bispectrum is defined as

$$\langle \tilde{\delta}(\mathbf{k}_1) \tilde{\delta}(\mathbf{k}_2) \tilde{\delta}(\mathbf{k}_3) \rangle_c = (2\pi)^3 \delta_D(\mathbf{k}_{123}) B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3). \quad (\text{A.15})$$

Since the connected bispectrum vanishes for Gaussian random fields, it is the first intrinsically non-linear moment. Inserting the perturbative expansion (A.4) for each term results generally in an infinitely large sequence of correlators. The lowest non-vanishing order is the so-called *tree-level* contribution to the bispectrum. We find for the correlator in tree level

$$\langle \tilde{\delta}(\mathbf{k}_1) \tilde{\delta}(\mathbf{k}_2) \tilde{\delta}(\mathbf{k}_3) \rangle_{\text{tree}} = \langle \tilde{\delta}_2(\mathbf{k}_1) \tilde{\delta}_1(\mathbf{k}_2) \tilde{\delta}_1(\mathbf{k}_3) \rangle + \langle \tilde{\delta}_1(\mathbf{k}_1) \tilde{\delta}_2(\mathbf{k}_2) \tilde{\delta}_1(\mathbf{k}_3) \rangle + \langle \tilde{\delta}_1(\mathbf{k}_1) \tilde{\delta}_1(\mathbf{k}_2) \tilde{\delta}_2(\mathbf{k}_3) \rangle. \quad (\text{A.16})$$

Replacing the second-order density contrast with Eq. (A.5) results in

$$\begin{aligned} \langle \tilde{\delta}_2(\mathbf{k}_1) \tilde{\delta}_1(\mathbf{k}_2) \tilde{\delta}_1(\mathbf{k}_3) \rangle &= \int \frac{d^3 q_1}{(2\pi)^3} \int d^3 q_2 \delta_D(\mathbf{k}_1 - \mathbf{q}_1 - \mathbf{q}_2) F_2(\mathbf{q}_1, \mathbf{q}_2) \langle \tilde{\delta}_1(\mathbf{q}_1) \tilde{\delta}_1(\mathbf{q}_2) \tilde{\delta}_1(\mathbf{k}_2) \tilde{\delta}_1(\mathbf{k}_3) \rangle \\ &= (2\pi)^3 \delta_D(\mathbf{k}_{123}) 2F_2(\mathbf{k}_2, \mathbf{k}_3) P_{\text{lin}}(k_2) P_{\text{lin}}(k_3), \end{aligned} \quad (\text{A.17})$$

where we applied Wick's theorem to express the four-point correlator of Gaussian fields in terms of products of power spectra, and performed the two integrations over the Dirac delta distributions. The results for the other two terms of the tree-level bispectrum are simply obtained by permutations of the arguments. Finally, the tree-level bispectrum is given by

$$B_{\text{pt}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 2F_2(\mathbf{k}_1, \mathbf{k}_2) P_1 P_2 + 2F_2(\mathbf{k}_1, \mathbf{k}_3) P_1 P_3 + 2F_2(\mathbf{k}_2, \mathbf{k}_3) P_2 P_3. \quad (\text{A.18})$$

The factor 2 follows from using the symmetrized version of the second-order coupling function.

### A.2.2. Trispectrum

The dark matter trispectrum is defined as the connected four-point function in Fourier space:

$$\langle \tilde{\delta}(\mathbf{k}_1) \tilde{\delta}(\mathbf{k}_2) \tilde{\delta}(\mathbf{k}_3) \tilde{\delta}(\mathbf{k}_4) \rangle_c = (2\pi)^3 \delta_D(\mathbf{k}_{1234}) T(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4), \quad (\text{A.19})$$

where  $\mathbf{k}_{1234} \equiv \mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 + \mathbf{k}_4$ . We find that there are two different non-vanishing contributions to the tree level:

$$\begin{aligned} \langle \tilde{\delta}(\mathbf{k}_1) \tilde{\delta}(\mathbf{k}_2) \tilde{\delta}(\mathbf{k}_3) \tilde{\delta}(\mathbf{k}_4) \rangle_{\text{tree}} &= \langle \tilde{\delta}_2(\mathbf{k}_1) \tilde{\delta}_2(\mathbf{k}_2) \tilde{\delta}_1(\mathbf{k}_3) \tilde{\delta}_1(\mathbf{k}_4) \rangle + \dots (6 \text{ terms}) \\ &+ \langle \tilde{\delta}_3(\mathbf{k}_1) \tilde{\delta}_1(\mathbf{k}_2) \tilde{\delta}_1(\mathbf{k}_3) \tilde{\delta}_1(\mathbf{k}_4) \rangle + \dots (4 \text{ terms}). \end{aligned} \quad (\text{A.20})$$

In total we find 6 terms of the first type and 4 terms for the second type, where the rest is obtained by permutations. All other contributions either vanish or are built of higher-order terms. Note that for the second type of terms we need the results from perturbation theory up to the third order. The calculation of each term is a tedious but straightforward calculation. We obtain for the first term of the expansion

$$\langle \tilde{\delta}_2(\mathbf{k}_1) \tilde{\delta}_2(\mathbf{k}_2) \tilde{\delta}_1(\mathbf{k}_3) \tilde{\delta}_1(\mathbf{k}_4) \rangle = (2\pi)^3 \delta_D(\mathbf{k}_{1234}) 4P_3 P_4 [P_{13} F_2(\mathbf{k}_3, -\mathbf{k}_{13}) F_2(\mathbf{k}_4, -\mathbf{k}_{24}) + P_{14} F_2(\mathbf{k}_3, -\mathbf{k}_{23}) F_2(\mathbf{k}_4, -\mathbf{k}_{14})], \quad (\text{A.21})$$

where the six-point correlator resolves into 15 terms consisting of power spectra products. Performing the integrations over the arising delta functions yields in the end 8 different terms. Similarly, we find for the second type of terms

$$\langle \tilde{\delta}_3(\mathbf{k}_1) \tilde{\delta}_1(\mathbf{k}_2) \tilde{\delta}_1(\mathbf{k}_3) \tilde{\delta}_1(\mathbf{k}_4) \rangle = (2\pi)^3 \delta_D(\mathbf{k}_{1234}) 6F_3(\mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) P_2 P_3 P_4. \quad (\text{A.22})$$

The other terms are easily obtained by permutations, however, we present here the complete result to avoid confusion with a shorthand notation that is introduced afterwards. The trispectrum of cold dark matter is in first non-vanishing order given by (Fry 1984):

$$T_{\text{pt}} = 4T_a + 6T_b, \quad (\text{A.23})$$

where

$$\begin{aligned} T_a &= P_1 P_2 [P_{13} F_2(\mathbf{k}_1, -\mathbf{k}_{13}) F_2(\mathbf{k}_2, \mathbf{k}_{13}) + P_{14} F_2(\mathbf{k}_1, -\mathbf{k}_{14}) F_2(\mathbf{k}_2, \mathbf{k}_{14})] \\ &+ P_1 P_3 [P_{12} F_2(\mathbf{k}_1, -\mathbf{k}_{12}) F_2(\mathbf{k}_3, \mathbf{k}_{12}) + P_{14} F_2(\mathbf{k}_1, -\mathbf{k}_{14}) F_2(\mathbf{k}_3, \mathbf{k}_{14})] \\ &+ P_1 P_4 [P_{12} F_2(\mathbf{k}_1, -\mathbf{k}_{12}) F_2(\mathbf{k}_4, \mathbf{k}_{12}) + P_{13} F_2(\mathbf{k}_1, -\mathbf{k}_{13}) F_2(\mathbf{k}_4, \mathbf{k}_{13})] \\ &+ P_2 P_3 [P_{21} F_2(\mathbf{k}_2, -\mathbf{k}_{21}) F_2(\mathbf{k}_3, \mathbf{k}_{21}) + P_{24} F_2(\mathbf{k}_2, -\mathbf{k}_{24}) F_2(\mathbf{k}_3, \mathbf{k}_{24})] \\ &+ P_2 P_4 [P_{21} F_2(\mathbf{k}_2, -\mathbf{k}_{21}) F_2(\mathbf{k}_4, \mathbf{k}_{21}) + P_{23} F_2(\mathbf{k}_2, -\mathbf{k}_{23}) F_2(\mathbf{k}_4, \mathbf{k}_{23})] \\ &+ P_3 P_4 [P_{31} F_2(\mathbf{k}_3, -\mathbf{k}_{31}) F_2(\mathbf{k}_4, \mathbf{k}_{31}) + P_{32} F_2(\mathbf{k}_3, -\mathbf{k}_{32}) F_2(\mathbf{k}_4, \mathbf{k}_{32})], \end{aligned} \quad (\text{A.24})$$

and

$$T_b = F_3(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) P_1 P_2 P_3 + F_3(\mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) P_2 P_3 P_4 + F_3(\mathbf{k}_3, \mathbf{k}_4, \mathbf{k}_1) P_3 P_4 P_1 + F_3(\mathbf{k}_4, \mathbf{k}_1, \mathbf{k}_2) P_4 P_1 P_2, \quad (\text{A.25})$$

where  $P_i \equiv P_{\text{lin}}(k_i)$ ,  $P_{ij} \equiv P_{\text{lin}}(|\mathbf{k}_i + \mathbf{k}_j|)$  and  $\mathbf{k}_{ij} \equiv \mathbf{k}_i + \mathbf{k}_j$ .

For the covariance matrix one only needs the parallelogram configuration. This imposes the condition  $\mathbf{k}_2 = -\mathbf{k}_1$  and  $\mathbf{k}_4 = -\mathbf{k}_3$  on the wave-vectors. In this case Eq. (A.23) simplifies to

$$\begin{aligned} T_{\text{pt}} = & 4P_1^2 \left\{ [F_2(\mathbf{k}_1, -\mathbf{k}_+)]^2 P_+ + [F_2(\mathbf{k}_1, \mathbf{k}_-)]^2 P_- \right\} + 4P_3^2 \left\{ [F_2(\mathbf{k}_3, -\mathbf{k}_+)]^2 P_+ + [F_2(\mathbf{k}_3, -\mathbf{k}_-)]^2 P_- \right\} \\ & + 8P_1 P_3 [F_2(\mathbf{k}_1, -\mathbf{k}_+) F_2(\mathbf{k}_3, -\mathbf{k}_+) P_+ + F_2(\mathbf{k}_1, \mathbf{k}_-) F_2(\mathbf{k}_3, -\mathbf{k}_-) P_-] \\ & + 12 \left[ P_1^2 P_3 F_3(\mathbf{k}_1, -\mathbf{k}_1, \mathbf{k}_3) + P_1 P_3^2 F_3(\mathbf{k}_1, \mathbf{k}_3, -\mathbf{k}_3) \right], \end{aligned} \quad (\text{A.26})$$

where  $\mathbf{k}_- \equiv \mathbf{k}_3 - \mathbf{k}_1$ ,  $\mathbf{k}_+ \equiv \mathbf{k}_1 + \mathbf{k}_3$ ,  $P_- \equiv P_{\text{lin}}(|\mathbf{k}_-|)$  and  $P_+ \equiv P_{\text{lin}}(|\mathbf{k}_+|)$ . Consequently, we define the non-Gaussian contribution to the covariance in tree-level perturbation theory as

$$C_{\text{pt}} \equiv C_{\text{pt}}(\ell_i, \ell_j) = \frac{1}{A} \int_{|\ell_i| \in \ell_i} \frac{d^2 \ell_1}{A_r(\ell_i)} \int_{|\ell_2| \in \ell_j} \frac{d^2 \ell_2}{A_r(\ell_j)} \frac{\ell_1^2 \ell_2^2}{(2\pi)^2} \int_0^{w_H} dw \frac{G^4(w)}{w^2} T_{\text{pt}} \left( \frac{\ell_1}{w}, -\frac{\ell_1}{w}, \frac{\ell_2}{w}, -\frac{\ell_2}{w}, w \right), \quad (\text{A.27})$$

where  $A$  denotes the survey area,  $A_r(l)$  the integration area and  $G(w)$  the lensing weight function (see Eq. 7). More details on the notation can be found in Sect. 3.

## References

- Albrecht, A., Bernstein, G., Cahn, R., et al. 2006, ArXiv:astro-ph/0609591  
 Bacon, D. J., Refregier, A. R., & Ellis, R. S. 2000, MNRAS, 318, 625  
 Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15  
 Bartelmann, M. & Schneider, P. 2001, Phys. Rep., 340, 291  
 Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002, Phys. Rep., 367, 1  
 Bullock, J. S., Kolatt, T. S., Sigad, Y., et al. 2001, MNRAS, 321, 559  
 Cooray, A. & Hu, W. 2001, ApJ, 554, 56  
 Cooray, A. & Sheth, R. 2002, Phys. Rep., 372, 1  
 Dolney, D., Jain, B., & Takada, M. 2004, MNRAS, 352, 1019  
 Eisenstein, D. J. & Hu, W. 1998, ApJ, 496, 605  
 Fry, J. N. 1984, ApJ, 279, 499  
 Fu, L., Semboloni, E., Hoekstra, H., et al. 2008, A&A, 479, 9  
 Gunn, J. E. & Gott, J. R. I. 1972, ApJ, 176, 1  
 Hartlap, J., Schrabback, T., Simon, P., & Schneider, P. 2009, ArXiv:0901.3269  
 Hilbert, S., Hartlap, J., White, S. D. M., & Schneider, P. 2008, ArXiv:0809.5035  
 Hu, W. & Kravtsov, A. V. 2003, ApJ, 584, 702  
 Jain, B. & Bertschinger, E. 1994, ApJ, 431, 495  
 Jain, B., Seljak, U., & White, S. 2000, ApJ, 530, 547  
 Jenkins, A., Frenk, C. S., Pearce, F. R., et al. 1998, ApJ, 499, 20  
 Jing, Y. P. 2000, ApJ, 535, 30  
 Joachimi, B., Schneider, P., & Eifler, T. 2008, A&A, 477, 43  
 Kaiser, N. 1992, ApJ, 388, 272  
 Kaiser, N. 1998, ApJ, 498, 26  
 Kaiser, N., Wilson, G., & Luppino, G. A. 2000, ArXiv:astro-ph/0003338  
 Kilbinger, M. & Schneider, P. 2005, A&A, 442, 69  
 Komatsu, E., Dunkley, J., Nolte, M. R., et al. 2009, ApJS, 180, 330  
 Ma, C.-P. & Fry, J. N. 2000, ApJ, 543, 503  
 Massey, R., Heymans, C., Bergé, J., et al. 2007, MNRAS, 376, 13  
 Nakamura, T. T. & Suto, Y. 1997, Progress of Theoretical Physics, 97, 49  
 Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493  
 Peacock, J. A. & Dodds, S. J. 1996, MNRAS, 280, L19  
 Pielorz, J. 2008, PhD thesis, University of Bonn  
 Press, W. H. & Schechter, P. 1974, ApJ, 187, 425  
 Sato, M., Hamana, T., Takahashi, R., et al. 2009, ArXiv:0906.2237  
 Schaefer, B. M. 2008, ArXiv:0808.0203  
 Schneider, P. 2006, in *Gravitational Lensing: Strong, Weak and Micro*, Saas-Fee Advanced Courses, Volume 33., ed. P. Schneider, C. S. Kochanek, & J. Wambsganss (Springer-Verlag, Berlin), 269  
 Schneider, P., van Waerbeke, L., Kilbinger, M., & Mellier, Y. 2002, A&A, 396, 1  
 Scoccimarro, R., Colombi, S., Fry, J. N., et al. 1998, ApJ, 496, 586  
 Scoccimarro, R., Sheth, R. K., Hui, L., & Jain, B. 2001, ApJ, 546, 20  
 Scoccimarro, R., Zaldarriaga, M., & Hui, L. 1999, ApJ, 527, 1  
 Seljak, U. 2000, MNRAS, 318, 203  
 Semboloni, E., Tereno, I., van Waerbeke, L., & Heymans, C. 2008, ArXiv:0812.1881  
 Semboloni, E., van Waerbeke, L., Heymans, C., et al. 2007, MNRAS, 375, L6  
 Sheth, R. K. & Tormen, G. 1999, MNRAS, 308, 119  
 Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, MNRAS, 341, 1311  
 Smith, R. E., Scoccimarro, R., & Sheth, R. K. 2007, Phys. Rev. D, 75, 063512  
 Smith, R. E. & Watts, P. I. R. 2005, MNRAS, 360, 203  
 Smith, R. E., Watts, P. I. R., & Sheth, R. K. 2006, MNRAS, 365, 214  
 Springel, V. 2005, MNRAS, 364, 1105  
 Takada, M. & Bridle, S. 2007, New Journal of Physics, 9, 446  
 Takada, M. & Jain, B. 2003, MNRAS, 340, 580  
 Takada, M. & Jain, B. 2009, MNRAS, 395, 2065  
 Takahashi, R., Yoshida, N., Takada, M., et al. 2009, ArXiv:0902.0371

- Tinker, J. L., Weinberg, D. H., Zheng, Z., & Zehavi, I. 2005, *ApJ*, 631, 41
- Van Waerbeke, L., Mellier, Y., Erben, T., et al. 2000, *A&A*, 358, 30
- Vikhlinin, A., Burenin, R. A., Ebeling, H., et al. 2009, *ApJ*, 692, 1033
- Wechsler, R. H., Bullock, J. S., Primack, J. R., Kravtsov, A. V., & Dekel, A. 2002, *ApJ*, 568, 52
- White, M. & Hu, W. 2000, *ApJ*, 537, 1
- Wittman, D. M., Tyson, J. A., Kirkman, D., Dell'Antonio, I., & Bernstein, G. 2000, *Nature*, 405, 143